ENUMERATIVE COMBINATORICS OF POSETS

A Thesis
Presented to
The Academic Faculty

by

Christina C. Carroll

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy in the
Department of Mathematics

Georgia Institute of Technology
April, 2008

ENUMERATIVE COMBINATORICS OF POSETS

Approved by:

Prasad Tetali,
(School of Mathematics and College of Computing), Advisor
Department of Mathematics
Georgia Institute of Technology

Richard Duke,
School of Mathematics, *emeritus*
Department of Mathematics
Georgia Institute of Technology

Christine Heitsch
Department of Mathematics
Georgia Institute of Technology

Dana Randall,
(College of Computing)
Department of Mathematics
Georgia Institute of Technology

William T. Trotter
Department of Mathematics
Georgia Institute of Technology

Date Approved: 14 March 2008

To David and Fern, both of whom at points almost derailed this project, but ultimately became the reason that it got done.

ACKNOWLEDGEMENTS

It takes a village to write a thesis. Although I cannot begin to list all of the people who have touched my life during this journey, there are several I want to make sure to thank by name:

Prasad: *For his attention to detail and for so many other things.*

Jamie, Carol, Patty, Tom, and Dana: *For mentoring which has touched this document in many ways, both large and small.*

Irene, Hamid, Sam, and Jennifer: *For teaching me how to stay sane during this process.*

Sam, Csaba, and Mitch: *For their support, technical and otherwise.*

Josh, Gyula, and David: *For their guidance and shared ideas.*

My family: *For their unwavering faith in me.*

TABLE OF CONTENTS

LIST OF FIGURES

CHAPTER I

INTRODUCTION AND BACKGROUND

In this thesis we present several problems in enumerative and asymptotic combinatorics. Though each component problem is distinct, the work explores several common themes. The fundamental structures for which we pose questions are primarily partially ordered sets. For a more in-depth discussion of the connections between the chapters, please refer to the Conclusion (Chapter 7).

Before describing the main contributions of this thesis, in Section 1.1 we will recall three classical results concerning posets and graphs, which provide the basis and motivation for most of our work. These are

(i) ***Sperner's theorem*** concerning the size of the largest antichain in the Boolean lattice,

(ii) ***Kleitman's theorem*** addressing the logarithm of the number of antichains in the Boolean lattice, and

(iii) ***Brégman's theorem*** concerning the maximal number of perfect matchings in bipartite graphs.

In order to state the classical theorems, as well as the results of this thesis, in the introduction, we assume knowledge of standard definitions concerning graphs and partially ordered sets. These definitions are given in Section 1.3; a reader unfamiliar with this area may want to consult this section before reading further. Section 1.4, gives background on the entropy function and how it is used; though this material is foundational for the rest of the thesis, it is not necessary for an understanding of the remainder of the introduction.

1.1 Classical Results

The following theorem was first proven by Sperner in 1928 [46] and has inspired many similar results in the study of combinatorics of finite sets. The following statement is from [16], a comprehensive and useful book focused on Sperner theory.

Theorem 1.1.1 *(Sperner's Theorem) Let n be a positive integer and $\mathcal{F}$ be a family of subsets of $[n] := \{1, \ldots, n\}$ such that no member of $\mathcal{F}$ is included in another, that is for all $X, Y \in \mathcal{F}$ we have $X \not\subset Y$. Then*

$$|\mathcal{F}| \leq \begin{cases} \binom{n}{\frac{n}{2}} & \textit{if } n \textit{ is even,} \\ \binom{n}{\frac{n-1}{2}} & \textit{if } n \textit{ is odd.} \end{cases} \tag{1}$$

Equality holds if and only if

$$\mathcal{F} \begin{cases} \{X \subset [n] : |X| = \frac{n}{2}\} & \textit{for } n \textit{ even,} \\ \{X \subset [n] : |X| = \frac{n-1}{2}\} \textit{ or } \{X \subset [n] : |X| = \frac{n-1}{2}\} & \textit{for } n \textit{ odd.} \end{cases} \tag{2}$$

Sperner's theorem observes that the largest *inclusion free family* in the Boolean Lattice, $\mathfrak{B}_n$, is the middle layer. A related problem, posed by Dedekind in 1897, asks 'How *many* inclusion free families are there in the Boolean Lattice?' Sperner's theorem gives us a convenient lower bound for this problem, as any subset of an inclusion free family is in itself inclusion free. In 1969, Kleitman first solved Dedekind's Problem asymptotically at the logarithmic level [30], proving the following theorem:

Theorem 1.1.2 *(Kleitman) Let $a(\mathfrak{B}_n)$ represent the number of inclusion free families in the Boolean Lattice. Then*

$$\log\left(a(\mathfrak{B}_n)\right) = \binom{n}{\lfloor \frac{n}{2} \rfloor}(1 + o(1)). \tag{3}$$

This theorem shows that the majority of inclusion free families in $\mathfrak{B}_n$ are indeed subsets of the middle layer. This basic result has been improved upon several times (see [29, 26]).

The final classical result we recall here has a slightly different flavor. It is a theorem of Brégman addressing Minc's conjecture and was originally posed in the context of permanents of 0-1 matrices. It has been reframed to address the number of perfect matchings in bipartite graphs:

Theorem 1.1.3 *(Brégman) Let G be a bipartite graph on N vertices with partition classes A and B. Suppose that the degree sequence of A is given by $\{r_i\}_{i=1}^{|A|}$. Let $\mathcal{M}_{\text{perfect}}(G)$ be the* *set of perfect matchings in G. Then:*

$$|\mathcal{M}_{\text{perfect}}(G)| \leq \prod_{i=1}^{|A|} (r_i!)^{\frac{1}{r_i}}.$$

It is a consequence of this theorem that when $2d$ divides N, the graph consisting of $\frac{N}{2d}$ disjoint copies of $K_{d,d}$ contains the greatest number of perfect matchings over all d-regular bipartite graphs on N vertices.

1.2 Results

The following problems are investigated within this thesis:

- In Chapter 2, we pose a question in the vein of Sperner's Theorem: 'What is the size of the largest family which contains no three distinct subsets satisfying $C \subset A \cap B$, $A \not\subset B$?' Our result in this chapter is distinguished from similar known results: to our knowledge it is the first result which finds the maximum size of family which excludes a substructure described by an **induced** property.

  Theorem 1.2.1 *(Carroll, Katona) Let $\mathcal{F}$ be a family of sets in $\mathfrak{B}_n$ satisfying the desired property. Then*

  $$\binom{n}{\lfloor \frac{n}{2} \rfloor}\left(1 + \frac{1}{n} + \Omega\left(\frac{1}{n^2}\right)\right) \leq \max(|\mathcal{F}|) \leq \binom{n}{\lfloor \frac{n}{2} \rfloor}\left(1 + \frac{2}{n} + O\left(\frac{1}{n^2}\right)\right).$$

- In Chapter 3, we address two conjectures. The first, posed by Friedland [19] is reminiscent of Brégman's theorem. Though it follows directly from the theorem that disjoint copies of $K_{d,d}$ contain the extremal number of *perfect* matchings, the theorem is not strong enough to give us information about the graph which contains the extremal number of matchings a fixed size. Friedland asked 'Given a d-regular graph G on N vertices, what is the maximum number of matchings of a fixed size which can be contained in G?'

  The second conjecture, made by Kahn [25], has a similar statement, but a different justification for making the conjecture. He asked: 'Given a d-regular graph G on N vertices, what is the maximum number of independent sets of a fixed size which can be contained in G?' He posed this conjecture after showing that disjoint copies of $K_{d,d}$ contain the maximum *total number* of independent sets.

For both of these problems, we use an entropy-based approach to provide an upper bound. Additionally, we show that $DK_{N,d}$, the graph consisting of $\frac{N}{2d}$ vertex-disjoint copies of the complete bipartite graph $K_{d,d}$, contains an asymptotically extremal number of matchings *(independent sets)* of a fixed size. Though the statements of these problems are very similar, the necessary techniques and strength of the results vary between the two solutions. For matchings, our result states that:

Theorem 1.2.2 *(Carroll, Galvin, Tetali) Let G be a d-regular graph on N vertices and ℓ an integer satisfying $0 \le \ell \le \frac{N}{2}$. Set $\alpha = \frac{2\ell}{N}$. Let $\mathcal{M}_\ell(G)$ represent the number of matchings in G of size ℓ. Then*

$$\log(\mathcal{M}_\ell(G)) \le \frac{N}{2}\left[\alpha \log d + H(\alpha)\right].$$

This bound is tight up to the first order term: for fixed $\alpha \in (0,1)$,

$$\log(\mathcal{M}_\ell(DK_{N,d})) \ge \frac{N}{2}\left[\alpha \log d + 2H(\alpha) + \alpha \log\left(\frac{\alpha}{e}\right) + \Omega\left(\frac{\log(d)}{d}\right)\right],$$

with the constant in the Ω term depending only on α.

The theorem about independent sets of a fixed size is more complicated to state in its entirety. Here we state the special case when $N = \omega(d \log d)$. See Section 3.1 for the full statement, results for different ranges of N and d, and a stronger result in the case that G is bipartite.

Theorem 1.2.3 *(Carroll, Galvin, Tetali) Let G be a d-regular graph on N vertices with $N = \omega(d \log d)$. Let ℓ be an integer satisfying $0 \le \ell \le \frac{N}{2}$. Set $\alpha = \frac{2\ell}{N}$. Let $i_\ell(G)$ represent the number of independent sets in G of size ℓ. Then*

$$\log(i_\ell(G)) \le \frac{N}{2}\left[H(\alpha) + \frac{2}{d}\right].$$

This bound is tight up to the second order term as :

$$\log(i_\ell(DK_{N,d})) \ge \frac{N}{2}\left[H(\alpha) + \frac{1}{d}(1 + o(1))\right].$$

(See Section 1.4 for a full discussion of the function $H(\alpha)$.)

- Given a partially ordered set, it is natural to define a total ordering of the vertices which is consistent with this partial ordering. Such a structure is called a linear extension of the poset. Over all bipartite posets on N vertices which have uniform up degrees u and down degrees d, Brightwell and Tetali [8] proved that the bipartite poset consisting of $\frac{N}{d+u}$ copies of $K_{d,u}$ has the most linear extensions. In Chapter 4, we generalize this result by providing a bound which can be applied on posets with up degrees bounded above by U and down degrees bounded below by D. We conclude the chapter by providing bounds for the number of linear extensions in $F_{n,k}$, the poset of partially defined functions. In this poset we order partially defined functions from $[n]$ to $[k]$ by inclusion. Given two such functions, $f \preceq g$ if and only if g is an extension of f. We can see that $\mathfrak{B}_n$ is isomorphic to $F_{n,1}$. See Section 4.2 for further introduction to this poset.

- Dedekind's problem is the inspiration for the problems in Chapter 5. We give results in a similar fashion to the logarithmic results of Kleitman addressing three questions. The first question is: How many antichains does $F_{n,k}$? We answer this at the logarithmic level showing that:

Theorem 1.2.4 *Let $a(P)$ represent the number of antichains in a poset P, and let $s_{n,k} = \left\lfloor \frac{kn}{(k+1)} \right\rfloor$. Then*

$$\binom{n}{s_{n,k}} k^{s_{n,k}} \leq \log(a(F_{n,k})) \leq \binom{n}{s_{n,k}} k^{s_{n,k}} \left(1 + O\left(\frac{\log(s_{n,k})}{s_{n,k}}\right)\right).$$

The second result of Chapter 5 addresses the number of antichains in $[t]^n$, the chain product poset.

Theorem 1.2.5 *(Carroll, Cooper, Tetali) Let for $n \geq 4$, $t = o(n^{\epsilon})$ for $0 < \epsilon \leq \frac{1}{8}$ and $\tau = t^n \sqrt{\frac{6}{\pi(t^2-1)n}}\,(1 + o(1))$. Then*

$$\tau \leq \log(a([t]^n)) \leq \tau \left(1 + \frac{11t^2(\log n)^{\frac{3}{2}}}{n^{\frac{1}{4}}}\right).$$

The poset $[t]^n$, though it is a generalization of $\mathfrak{B}_n$, is not regular. The differing degrees within level sets makes giving estimates on the number of antichains more difficult. We extend the method of Pippenger [36] to overcome the issue of varying degrees within levels.

We conclude the chapter by introducing the structure of a "butterfly" and giving bounds on how many butterfly-free families there are in the Boolean Lattice. See Section 5.3 for this result.

- In the final technical chapter we provide another Dedekind type result, though the structure we forbid is more complicated than inclusion-free. We bound the number of families in bipartite regular posets which exclude the 'cherry' structure; three sets A, B, C, where $A \prec C$ and $B \prec C$. We use an entropy-based upper bound and an algorithmic lower bound to show the following theorem:

  Theorem 1.2.6 *Let P be a regular two-level poset with n points on top with down degree d and m points on the bottom with up degree u. Then, letting $\mathfrak{C}(P)$ represent the number of cherry-free families on a poset P, we have*

$$n + \frac{m}{ud - u + 1} \leq \log\left(\mathfrak{C}(P)\right) \leq n + \frac{m}{d}\log(d+1).$$

  Moreover, this upper bound is tight up to constants in the second order term, as it is achieved by $\frac{n}{d}$ disjoint copies of $K_{d,d}$.

  Additionally, we provide an improved lower bound in the case that P is C_4-free. At the end of the chapter, we conclude by applying our lower bound to estimate the number of Horn functions in the Boolean lattice and compare this bound to known results.

We conclude the thesis with a summary and a description of further research directions.

1.3 Definitions

Here we provide a glossary of terms having to do with graphs and partially ordered sets. A reader familiar with standard definitions can safely skip to the next section.

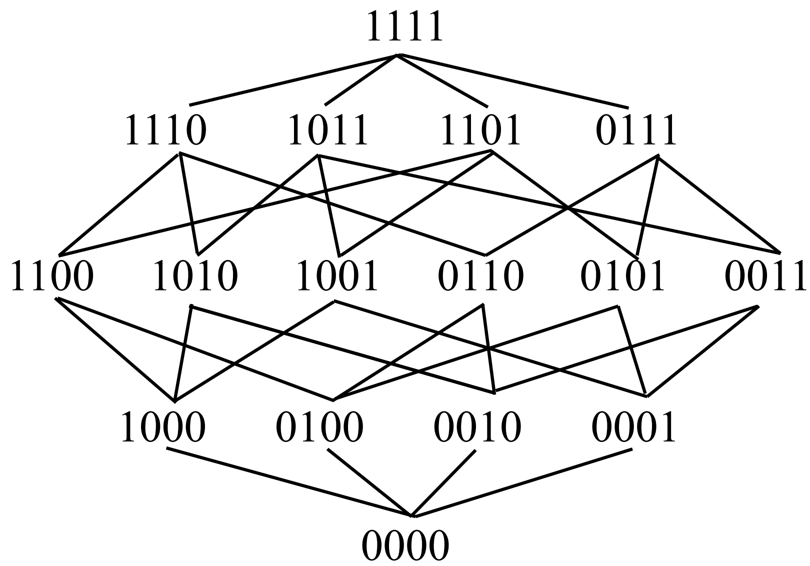

Figure 1: A classic poset: $\mathfrak{B}_4$

Definition 1.3.1 *A **graph** is a set of elements V, called **vertices**, and E, a collection of pairs of vertices, called **edges**. We usually insist that an edge consists of two distinct vertices. We use the notation $G(V, E)$, but when the vertex and edge sets are clear we will refer merely to G. The **degree of a vertex** is simply the number of edges in which the vertex is included.*

We define a special class of graphs with additional structure:

Definition 1.3.2 *A **partially ordered set** (also referred to as a **poset**) is a pair consisting of a ground set and a relation. The relation is reflexive, transitive and anti-symmetric. We use the notation $P(X, \preceq)$ where X is the ground set and $\preceq$ is the relation. If both X and $\preceq$ are clear from context, we will suppress the notation, referring merely to P.*

Definition 1.3.3 *The n-dimensional **Boolean Lattice** is the poset consisting of all subsets of $[n]$ ordered by inclusion. We use the notation $\mathfrak{B}_n$ to represent this poset.*

Definition 1.3.4 *In a graph, an **independent set** is a collection of vertices, no two of which are pairwise adjacent; i.e., at most one endpoint of each edge can be included in the* *independent set.*

In the context of posets, we can strengthen the idea of an independent set:

Definition 1.3.5 *In a poset $P(X, \preceq)$, an **antichain** is a collection of elements no two of which are related under $\preceq$.*

For example, an antichain $\mathcal{A}$ in the Boolean Lattice is a family of subsets of $[n]$ which satisfies that if A and B are distinct elements of $\mathcal{A}$ then $A \not\subset B$.

Definition 1.3.6 *A **matching** $\mathcal{M}$ is a collection of edges in a graph $G(V, E)$ so that each vertex in V is incident to at most one of the edges included in $\mathcal{M}$. If a matching $\mathcal{N}$ satisfies that every vertex is incident to exactly one edge in $\mathcal{N}$, then we say that it is a **perfect matching**.*

Definition 1.3.7 *Let $P(X, \preceq)$ be a poset. For distinct $x \preceq z$ with no y so that $x \preceq y \preceq z$, we say that z **covers** x in the poset. We use the notation $x \lessdot z$. A drawing of the poset which shows vertices and their covering relations is called a **Hasse diagram**.*

Taking the transitive closure of the covering relations shown in a Hasse diagram recovers all of the relationship information needed to reconstruct the poset. Hasse diagrams give us a visual way of representing posets; see Figures 1.3 and 4.2 for examples.

Definition 1.3.8 *A poset P is said to be **ranked** if there is a rank function r defined on the set of vertices and if $x \lessdot y$, then $r(x) + 1 = r(y)$. We refer to elements of the same rank as a **level** or **level set** of the poset. If K is the maximum rank in P, we say that P has **height** K.*

Definition 1.3.9 *A ranked poset is considered **graded** if every minimal element has rank 0 and and every maximal element has the same rank.*

Definition 1.3.10 *A graph is called **regular** if every vertex has the same degree. A poset is called **regular** if it is graded and the number of upward and downward neighbors each element has is completely determined by its rank. The number of neighbors of an element*

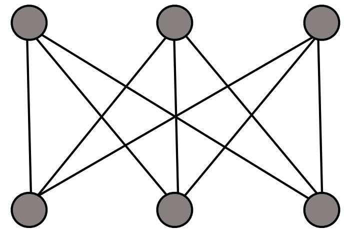

Figure 2: $K_{3,3}$: An example of a complete bipartite graph.

is referred to as its ***up degree*** *and* ***down degree*** *respectively. Given a vertex x we use the notation $d_{up}(x)$ and $d_{down}(x)$ for its respective degrees.*

Definition 1.3.11 *A graph is* ***bipartite*** *if it can be partitioned into two sets so that no element is adjacent to any other element in the same set.*

A very special graph (which can also be thought of as a poset) is the complete bipartite graph of degree d, referred to as $K_{d,d}$. It consists of $2d$ vertices partitioned into A and B, two classes of the same size. Every element in A is connected to every element in B, and there are no edges present within each partition class.

Definition 1.3.12 *A poset is called* ***Sperner*** *if the largest antichain and the largest level set have the same cardinality.*

(See [16, 48] for additional information and definitions concerning posets, and [3, 15] for additional graph theory background.)

1.4 Entropy as an Enumeration Tool

Entropy is a concept which has its origins in physics. The word calls up images of the competition between order and chaos, indeed the second law of thermodynamics states that "The entropy of an isolated system not in equilibrium will tend to increase over time, approaching a maximum value at equilibrium." The word gained its information theoretical meaning when Claude Shannon used it to describe the rate of a code. Although the contexts seem quite different, entropy is a single measure which can describe the degree of disorder in disparate situations.

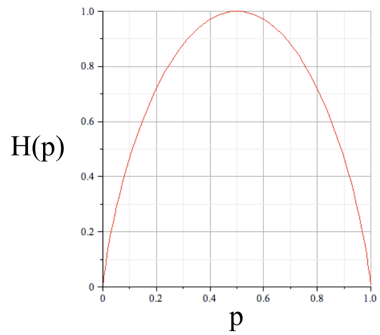

Figure 3: A graph of the binary entropy function

The entropy of a random variable has many desirable properties which allow us to manipulate it to get upper bounds on how much expected information the variable contains. If the outcome of a random event has high certainty the entropy of the random variable will be low; therefore we can think of entropy as proportional to uncertainty. This introduction will describe properties of entropy which will be used in the rest of the document.

Definition 1.4.1 *Let X be a finitely valued, discrete random variable taking values $\{x_i\}_{i=0}^n$. Let $p_i = \mathbb{P}(X = x_i)$. Then the **entropy** of X is*

$$H(X) := \sum_{x_i} p_i \log_2 \left(\frac{1}{p_i} \right). \tag{4}$$

Unless otherwise noted, $\log x$ will always represent $\log_2 x$. When X is a Bernoulli random variable, with probability p of success, $H(X)$ is known as the binary entropy function and we use the simplified notation $H(p)$. Furthermore, for any $p \in (0, 1)$ we define $H(p) = -p \log p - (1-p) \log(1-p)$, with $H(0) = H(1) = 0$ for continuity (see figure 1.4).

Proofs of the elementary properties of entropy can be found in many sources. The proofs below were adapted from those found in [2]. We will frequently use the concavity of the function $\log(x)$, primarily because it allows us to employ Jensen's inequality.

Proposition 1.4.2 *(Jensen's Inequality) Let $\phi(x)$ be a concave function with values $\{x_i\}_{i=1}^n$ in its domain. For each x_i, assign a nonnegative weight p_i, so that $\sum_{i=1}^n p_i = 1$. Then following property holds:*

$$\phi\left(\sum_i^n p_i x_i \right) \geq \sum_i^n p_i \phi(x_i).$$

This holds with equality if and only if all p_i are equal.

The proof for Jensen's inequality will not be provided here, but it follows directly from the definition of concavity and a simple induction argument.

Definition 1.4.3 *Suppose X is a random variable which takes values $\{x_i\}_{i=1}^N$ with associated probabilities $\{p_i\}_{i=1}^N$. Similarly let Y be a random variable taking values $\{y_j\}_{j=1}^M$ with probabilities $\{q_j\}_{j=1}^M$. We define the* ***joint distribution*** *of X and Y as the random variable (X, Y) taking values of the form (x_i, y_j) for $1 \leq i \leq N$ and $1 \leq j \leq M$ with probabilities $r_{ij} := \mathbb{P}(X = x_i, Y = y_j)$.*

Proposition 1.4.4 *Let X and Y be two random variables, then*

$$H(X,Y) \geq H(X).$$

Proof. We expand $H(X,Y)$ using the definition of entropy and then using the fact that for all i, j, $r_{ij} \leq p_i$, we have:

$$\begin{aligned} H(X,Y) &= \sum_{i=1}^{N}\sum_{j=1}^{M} r_{ij} \log\left(\frac{1}{r_{ij}}\right) \\ &\geq \sum_{i=1}^{N}\sum_{j=1}^{M} r_{ij} \log\left(\frac{1}{p_i}\right) \\ &= \sum_{i=1}^{N} p_i \log\left(\frac{1}{p_i}\right) \\ &= H(X). \end{aligned} \tag{5}$$

Proposition 1.4.5 *If a random variable has support of size N, then*

$$H(X) \leq \log N,$$

with equality if and only if X is a uniform random variable.

Proof. We use Jensen's Inequality with concave function $\phi(x) = \log(x)$.

$$\begin{aligned} H(X) &= \sum_{i=1}^{N} p_i \log\left(\frac{1}{p_i}\right) \\ &\leq \log\left(\sum_{i=1}^{N} p_i \frac{1}{p_i}\right) \\ &= \log N. \end{aligned} \tag{6}$$

We can see that if X is a uniform random variable, (i.e. $p_i = \frac{1}{N}$ for all i) then equality is achieved. The other direction of the if and only if statement in this proposition follows from the equality conditions of Jensen's inequality.

Definition 1.4.6 *Using the definitions for X and Y as above, the **conditional entropy** of X given Y denoted $H(X|Y)$, is defined to be*

$$\begin{aligned} H(X|Y) &= \sum_{j=1}^{M} \mathbb{P}(Y = y_j) H(X|Y = y_j) \\ &= -\sum_{j=1}^{M} \mathbb{P}(Y = y_j) \sum_{i=1}^{N} \mathbb{P}(Y = y_j|X = x_i) \log(\mathbb{P}(Y = y_j|X = x_i)). \end{aligned} \tag{7}$$

Proposition 1.4.7 *(Chain Rule for conditional entropy)*

$$H(X, Y) = H(X) + H(Y|X).$$

The proof follows directly from the definitions and Bayes' Theorem, so will be omitted.

Proposition 1.4.8 *(Subadditivity of Entropy). If X is a vector of random variables, $X = (X_1, X_2, \ldots, X_n)$, then*

$$H(X) \leq \sum_{i=1}^{n} H(X_i).$$

Proof. This result follows from the two variable case $H(X, Y) \leq H(X) + H(Y)$ and a straightforward induction proof. We will provide a proof of the base case by showing that $H(X) + H(Y) - H(X, Y) \geq 0$. To begin, we expand each of these terms using the r_{ij} notation introduced in the definition of joint distribution above.

$$\begin{aligned} &H(X) + H(Y) - H(X, Y) \\ &= -\sum_{i=1}^{N} \mathbb{P}(X = x_i) \log(\mathbb{P}(X = x_i)) - \sum_{j=1}^{M} \mathbb{P}(Y = y_j) \log(\mathbb{P}(Y = y_j)) + \sum_{i=1}^{N} \sum_{j=1}^{M} r_{ij} \log\left(r_{ij}\right) \\ &= \sum_{i=1}^{N} \sum_{j=1}^{M} r_{ij} \log\left(\frac{r_{ij}}{\mathbb{P}(X = x_i)\mathbb{P}(Y = y_i)}\right) \\ &= \sum_{i=1}^{N} \sum_{j=1}^{M} \mathbb{P}(X = x_i)\mathbb{P}(Y = y_i) \frac{r_{ij}}{\mathbb{P}(X = x_i)\mathbb{P}(Y = y_i)} \log\left(\frac{r_{ij}}{\mathbb{P}(X = x_i)\mathbb{P}(Y = y_i)}\right). \end{aligned} \tag{8}$$

We have put the final terms of this sum into the form $z_{ij} \log z_{ij}$ with $z_{ij} = \frac{r_{ij}}{\mathbb{P}(X=x_i)\mathbb{P}(Y=y_i)}$. Since the function $f(z) = z \log z$ is convex, we can use Jensen's Inequality to yield:

$$H(X) + H(Y) - H(X,Y) \geq f\left(\sum_{i=1}^{N}\sum_{j=1}^{M} \mathbb{P}(X = x_i)\mathbb{P}(Y = y_i) z_{ij}\right) = f(1) = 0.$$

Definition 1.4.9 *Let $X = (X_1, X_2, \ldots, X_n)$ be a random variable. Let $\mathcal{G}$ be a family of subsets of $[n]$. We call $\mathcal{G}$ a* ***covering family*** *of degree k if each $i \in [n]$ belongs to at least k members of $\mathcal{G}$.*

There are many useful variations of the following lemma originally proven by Shearer [11]; In particular it was pointed out in [33] that this result is implicit in the work of game theorists Bondareva [5] and Shapely [41, 42]. See [18] for a weighted version, [33] for a fractionally subadditive version of Shearer's Lemma.

We will present only the most basic version here. For a subset $A \subseteq [n]$, we will use the notation X_A for $(X_i)_{i \in A}$.

Lemma 1.4.10 *(Shearer's Lemma) Let $X = (X_1, X_2, \ldots, X_n)$ be a random variable. If $\mathcal{G}$ is a covering family of degree k, then:*

$$kH(X) \leq \sum_{G \in \mathcal{G}} H(X_G).$$

Proof. (The proof we give was first given by Radhakrishnan [37].) Using the chain rule repeatedly we can rewrite $H(X)$ as a sum of conditional entropies:

$$H(X) = \sum_{j=1}^{n} H\big(X_j | (X_\ell : \ell < j)\big). \tag{9}$$

Similarly for $G \in \mathcal{G}$,

$$H(X_G) = \sum_{j \in G} H\big(X_j | (X_\ell : \ell \in G, \ell < j)\big). \tag{10}$$

Starting with the right hand side from the proposition, we note that for each $H(X_G)$ we can drop some of the conditioning, then change the order of summation, yielding the calculations:

$$
\begin{aligned}
\sum_{G\in\mathcal{G}} H(X_G) &= \sum_{G\in\mathcal{G}}\sum_{j\in G} H\big(X_j|(X_\ell : \ell \in G, \ell < j)\big) \\
&\geq \sum_{G\in\mathcal{G}}\sum_{j\in G} H\big(X_j|(X_\ell : \ell < j)\big) \\
&= \sum_{j\in[n]}\sum_{G\in\mathcal{G}:j\in G} H\big(X_j|(X_\ell : \ell < j)\big) \qquad (11)\\
&\text{(Since } \mathcal{G} \text{ is a covering family of degree } k\text{),} \\
&\geq k\sum_{j=1}^{n} H\big(X_j|(X_\ell : \ell < j)\big) \\
&= kH(X)
\end{aligned}
$$

Proposition 1.4.11 $H'(x) = \log\left(\frac{1-x}{x}\right)$.

Proof. The proof here is not mysterious; we use the product rule:

$$
\begin{aligned}
H(x) &= -x\log x - (1-x)\log(1-x) \\
H'(x) &= -\frac{x}{x} - \log x + \frac{1-x}{1-x} + \log(1-x) \qquad (12)\\
&= \log\left(\frac{1-x}{x}\right).
\end{aligned}
$$

CHAPTER II

NO THREE DISTINCT SUBSETS SATISFYING $C \subset A \cap B, A \not\subset B$

2.1 *Introduction*

Let $[n] = \{1, 2, \ldots, n\}$ be a finite set and $\mathcal{F} \subset 2^{[n]}$ a family of its subsets. In the present chapter, $\max |\mathcal{F}|$ will be investigated when $\mathcal{F}$ contains no three distinct subsets satisfying $C \subset A \cap B, A \not\subset B$. The well-known Sperner's Theorem ([46]) was the first such discovery of bounding the maximum size of a family which excludes certain substructures:

Theorem 2.1.1 *(Sperner's Theorem) If $\mathcal{F}$ is a family of subsets of $[n]$ without inclusion ($F, G \in \mathcal{F}$ implies $F \not\subset G$) then*

$$|\mathcal{F}| \leq \binom{n}{\lfloor \frac{n}{2} \rfloor}$$

holds, and this estimate is sharp as the family of all $\lfloor \frac{n}{2} \rfloor$-element subsets achieves this size.

There are a very large number of generalizations and analogues of this theorem. Here we will mention only some results where the conditions on $\mathcal{F}$ exclude certain configurations which can be expressed by inclusion only (i.e. the conditions can be stated without using intersections, unions, etc.) The first such generalization was obtained by Erdős [17].

Definition 2.1.1 *The family of k distinct sets with mutual inclusions, $F_1 \subset F_2 \subset \ldots F_k$ is called a* **chain of length** *k, which we denote simply by P_k.*

For any "small" family of sets $\mathcal{P}$, with specified inclusions between pairs of sets, let $\mathrm{La}(n, \mathcal{P})$ denote the size of the largest family $\mathcal{F}$ of subsets of $[n]$ which contains no $\mathcal{P}$ as a subfamily. In the rest of the chapter, the specified $\mathcal{P}$s will be denoted by normal upper case letters. Erdős extended Sperner's Theorem as follows:

Theorem 2.1.2 *(Erdős* [17]) $\mathrm{La}(n, P_{k+1})$ *is equal to the sum of the k largest binomial coefficients of order n. This bound is tight as the middle k layers of the Boolean lattice form a family of this size which contains no P_{k+1}.*

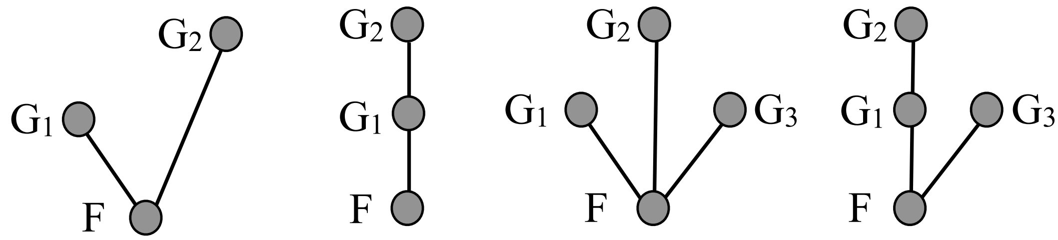

Figure 4: Possible configurations for V_2 and V_3 respectively

We can also consider families other than chains.

Definition 2.1.2 *An* **r-fork** *is a family of* $r+1$ *distinct sets so that* $F \subset G_1, F \subset G_2, \ldots, F \subset G_r$. *We use the notation* V_r *for an r-fork. See Figure 2.1.*

Notice that the class of 2-forks includes paths on 3 vertices; Similarly for r-forks we cannot assume anything about the relationships between the sets $\{G_i\}_{i=1}^r$. The quantity $\mathrm{La}(n, V_r)$ was first (asymptotically) determined for $r = 2$.

Theorem 2.1.3 *(Katona, Tarján* [28]*)*

$$\binom{n}{\lfloor \frac{n}{2} \rfloor}\left(1 + \frac{1}{n} + \Omega\left(\frac{1}{n^2}\right)\right) \leq \mathrm{La}(n, V_2) \leq \binom{n}{\lfloor \frac{n}{2} \rfloor}\left(1 + \frac{2}{n}\right).$$

This was recently generalized to forks with more branches by DeBonis *et. al* [12], as well as independently by Tranh [47]:

Theorem 2.1.4 *(DeBonis, Katona; Tranh)*

$$\binom{n}{\lfloor \frac{n}{2} \rfloor}\left(1 + \frac{r}{n} + \Omega\left(\frac{1}{n^2}\right)\right) \leq \mathrm{La}(n, V_{r+1}) \leq \binom{n}{\lfloor \frac{n}{2} \rfloor}\left(1 + 2\frac{r}{n} + O\left(\frac{1}{n^2}\right)\right).$$

The family of four distinct subsets satisfying $A \subset C, A \subset D, B \subset C$ is called and denoted by N. Another recent result is the following one:

Theorem 2.1.5 *(Griggs, Katona* [22]*)*

$$\binom{n}{\lfloor \frac{n}{2} \rfloor}\left(1 + \frac{1}{n} + \Omega\left(\frac{1}{n^2}\right)\right) \leq La(n, N) \leq \binom{n}{\lfloor \frac{n}{2} \rfloor}\left(1 + \frac{2}{n} + O\left(\frac{1}{n^2}\right)\right)$$

holds.

The goal of the present chapter is to investigate what happens if V_2 is excluded in an "induced way" that is only when the two "upper" sets are not related by inclusion. In other words, V_2 is excluded unless it is a P_3. Let $\mathrm{La}^\sharp(n, V_2)$ denote the size of the largest family $\mathcal{F}$ of subsets of $[n]$ containing no three distinct members $F, G_1, G_2 \in \mathcal{F}$ such that $F \subset G_1, F \subset G_2, G_1 \not\subset G_2$. We will refer to such a family as an *induced V-free* family. We prove the following sharpening of Theorem 2.1.3:

Theorem 2.1.6

$$\binom{n}{\lfloor \frac{n}{2} \rfloor}\left(1+\frac{1}{n}+\Omega\left(\frac{1}{n^2}\right)\right) \le \mathrm{La}^\sharp(n, V_2) \le \binom{n}{\lfloor \frac{n}{2} \rfloor}\left(1+\frac{2}{n}+O\left(\frac{1}{n^2}\right)\right).$$

Since $\mathrm{La}(n, V_2) \le \mathrm{La}^\sharp(n, V_2)$, the lower estimate is a consequence of the lower estimate in Theorem 2.1.3, we have to prove only the upper estimate. Although it seems to be a small modification, the proof (at least the one we found) is much more difficult than the proof of the upper estimate in Theorem 2.1.3. We will point out later what the differences are and why this case is more difficult. The method of the proof is a further extension of the proof used in [22].

The reader might be puzzled by the origin of the factors 1 and 2 in the respective second terms from the lower and upper estimates in Theorems 2.1.3 and 2.1.6. The construction proving the lower estimates (see [28] and [22]) is based on choosing the largest possible family $A_1, \dots, A_m$ of $\lceil \frac{n+1}{2} \rceil$-element sets such that $|A_i \cap A_j| < \lceil \frac{n-1}{2} \rceil$ $(i \neq j)$. The best known construction, given by Graham and Sloane (see [21]), gives only about half of the trivial upper estimate. See Section 2.5 for this construction. Finding the largest such set system is equivalent to a well known open problem of coding theory: what is the size of the largest binary code of length n with constant weight and minimum distance at least 4? The problem will be reduced to middle sized sets in Section 2.2.

Section 2.3 will give a sketch of the main idea of the proof, with details given in Section 2.4. We will give a version of Graham and Sloane's proof of lower bound in Section 2.5. We will conclude the chapter with a section giving a probabilistic view of the proof.

Note that because of the symmetry of the Boolean lattice and reflection around the middle layer, this chapter also proves the analogous result for the size of Boolean families which contain no three distinct sets $A \cup B \subset C, A \not\subset B$.

2.2 Reduction to middle sized sets

Observe that the main part of a large family must be near the middle since the total number of sets far from the middle is small. More precisely, let $0 < \alpha < \frac{1}{2}$ be a fixed real number. The total number of sets F (for a given n) of size satisfying

$$|F| \notin \left[n\left(\frac{1}{2} - \alpha\right), n\left(\frac{1}{2} + \alpha\right) \right] \tag{13}$$

is very small. It is well-known (see e.g. [2], page 232) that for a fixed constant $0 < \beta < \frac{1}{2}$

$$\sum_{i=0}^{\beta n} \binom{n}{i} = 2^{n(h(\beta)+o(1))}$$

holds where $h(x)$ is the binary entropy function: $h(x) = -x \log_2 x - (1-x)\log_2(1-x)$. Therefore using symmetry and the above fact, the total number of sets satisfying (13) is at most

$$2 \sum_{i=0}^{\lfloor n\left(\frac{1}{2}-\alpha\right)\rfloor} \binom{n}{i} \leq 2^{n\left(h\left(\frac{1}{2}-\alpha\right)+o(1)\right)} = \binom{n}{\lfloor \frac{n}{2} \rfloor} O\left(\frac{1}{n^2}\right) \tag{14}$$

where $0 < h\left(\frac{1}{2} - \alpha\right) < 1$ is a constant. We will prove the following theorem in Section 2.4.

Theorem 2.2.1 *If $\mathcal{F}$ satisfies the conditions of Theorem 2.1.6 and all members $F \in \mathcal{F}$ satisfy*

$$n\left(\frac{1}{2} - \alpha\right) \leq |F| \leq n\left(\frac{1}{2} + \alpha\right) \tag{15}$$

then

$$|\mathcal{F}| \leq \binom{n}{\lfloor \frac{n}{2} \rfloor} \left(1 + \frac{2}{n(1-2\alpha)} + O\left(\frac{1}{n^2}\right)\right). \tag{16}$$

Let us show that Theorem 2.2.1 implies Theorem 2.1.6.

If $\mathcal{F}$ is a family of distinct subsets of $[n]$, $0 < \alpha < \frac{1}{2}$ then let $\mathcal{F}_\alpha$ denote the subfamily consisting of sets satisfying (15). On the other hand, let $\mathcal{F}_{\overline{\alpha}}$ denote $\mathcal{F} - \mathcal{F}_\alpha$. If $\mathcal{F}$ satisfies the conditions of Theorem 2.1.6, then, by Theorem 2.2.1, (16) gives an upper estimate on $|\mathcal{F}_\alpha|$. On the other hand

$$|\mathcal{F}_{\overline{\alpha}}| \leq \binom{n}{\lfloor \frac{n}{2} \rfloor} O\left(\frac{1}{n^2}\right)$$

holds by (14). Since $|\mathcal{F}| = |\mathcal{F}_\alpha| + |\mathcal{F}_{\overline{\alpha}}|$, (16) and (2.5) imply that (16) is true not only for $\mathcal{F}_\alpha$, but also for $\mathcal{F}$, itself. Since we can write $|\mathcal{F}| = |\mathcal{F}_\alpha| + |\mathcal{F}_{\overline{\alpha}}|$, for any α, we see that $|\mathcal{F}|$ meets the bound in Theorem 2.1.6. Since the bound holds for every positive α, we see that it must hold for every $\mathcal{F}$, meaning that it must hold when $\alpha = 0$ as well, thus proving the Theorem.

2.3 *Plan of the proof*

We start off with some definitions:

Definition 2.3.1 *A family $\mathcal{G}$ is* **connected** *if for any pair (G_0, G_k) of its members there is a sequence $G_1, \ldots, G_{k-1}$ $(G_i \in \mathcal{G})$ such that either $G_i \subset G_{i+1}$ or $G_i \supset G_{i+1}$ holds for $0 \leq i < k$.*

Definition 2.3.2 *If a family is not connected, maximal connected subfamilies are called its* **connected components***.*

Definition 2.3.3 *A* **full chain** *in $2^{[n]}$ is a family of sets $A_0 \subset A_1 \subset \ldots \subset A_n$ where $|A_i| = i$.*

Note that that the number of full chains in $2^{[n]}$ is $n!$ since a full chain includes an additional one of the n elements from the ground set each level it passes through. Each of $n!$ permutations of $[n]$ yields a distinct full chain.

We say that a (full) chain $\mathcal{C}$ goes through a family $\mathcal{F}$ if they intersect, that is, if $\mathcal{F} \cap \mathcal{C} \neq \emptyset$. Let $\mathcal{F}_1, \ldots, \mathcal{F}_K$ be the components of $\mathcal{F}$, satisfying the conditions of Theorem 2.2.1. Denote the number of full chains going through $\mathcal{F}_i$ by $c(\mathcal{F}_i)$. Observe that a full chain cannot go through two distinct components since this would force them to be connected. Therefore, the following inequality holds:

$$\sum_{i=1}^{K} c(\mathcal{F}_i) \leq n!. \tag{17}$$

What can these components be? It is obvious that they have a tree-like form: each $\mathcal{F}_i$ has a maximal member which can contain several members, each of which can contain other members, and so on. For comparison, let us mention that in the case of the Sperner theorem, each component consists of exactly one set. In the case of Theorem 2.1.3, a component has a maximal member which contains an unlimited number of other members but they in turn cannot contain additional members; the longest chain that such a structure can contain has length two. In the present case, not only is the number of members unlimited so is the height of the longest included chain; this makes the proof more difficult.

We will give a good lower estimate

$$f(n,\alpha) \leq \frac{c(\mathcal{F}_i)}{|\mathcal{F}_i|}, \tag{18}$$

which will hold for all components $\mathcal{F}_i$. Together, (17) and (18) imply

$$f(n,\alpha)\sum_{i=1}^{K}|\mathcal{F}_i| \leq \sum_{i=1}^{K}\frac{c(\mathcal{F}_i)}{|\mathcal{F}_i|}|\mathcal{F}_i| = \sum_{i=1}^{K}c(\mathcal{F}_i) \leq n!.$$

Hence the final result will be

$$|\mathcal{F}| = \sum_{i=1}^{K}|\mathcal{F}_i| \leq \frac{n!}{f(n,\alpha)}, \tag{19}$$

which will prove Theorem 2.2.1 using the appropriate $f(n,\alpha)$.

The proof of the lower estimate (18) is based on the principle of inclusion/exclusion sieve, more precisely on a very primitive version. The number of chains going through $\mathcal{F}_i$ can be lower bounded by the sum of the number of chains going through the members $F \in \mathcal{F}_i$ minus the sum of the chains going through two members $F, G \in \mathcal{F}_i$ where F and G are comparable. This sum can be partitioned into $|\mathcal{F}_i|$ sums, where one sum consists of the number of chains going through a fixed $F \in \mathcal{F}_i$ minus the sum (over G) of the number of chains going through F and another member G such that $F \subset G$. If this sum is lower bounded by $f(n,\alpha)$, it implies (18). The proof of this latter estimate uses two facts: (i) For a given F, there is at most one set G with $F \subset G$ on each level. (ii) There are at most $2n\alpha$ such sets G. (i) is obtained from the condition of Theorem 2.1.6, while (ii) is a consequence of the condition (15).

Let us note that the restriction (15) was introduced because the estimate obtained from the first two terms of the sieve is too weak when either small or large sets are present in the family.

2.4 Details of the proof

We will follow the outline of the proof from the previous section, but prove the claims in reverse order.

Let $c(F)$ denote the number of full chains going through the set F and $c(F, G)$ the number of full chains going through both F and G. (This is obviously 0 if the two sets are incomparable.) If $\mathcal{G}$ is a family of subsets with $F \in \mathcal{G}$, define $d(\mathcal{G}, F)$ by

$$d(\mathcal{G}, F) = c(F) - \sum_{G \in \mathcal{G}:\ F \subset G} c(F, G).$$

Lemma 2.4.1 *Suppose $0 < \alpha \leq \frac{1}{8}$ and $n \geq 16$. Let $\mathcal{F}_i$ be a connected component of a family satisfying the conditions of Theorem 2.2.1. Then*

$$\frac{n!}{\binom{n}{\lfloor \frac{n}{2} \rfloor}} \left(1 - \frac{2}{n(1-2\alpha)} - \frac{128}{n^2}\right) \leq d(\mathcal{F}_i, F)$$

holds for every component F_i.

Proof.

$$\begin{aligned} d(\mathcal{F}_i, F) &= c(F) - \sum_{G \in \mathcal{F}_i:\ F \subset G} c(F, G) \\ &= |F|!(n - |F|)! - \sum_{G \in \mathcal{F}_i:\ F \subset G} |F|!(|G| - |F|)!(n - |G|)! \\ &= \frac{n!}{\binom{n}{|F|}} \left(1 - \sum_{G \in \mathcal{F}_i:\ F \subset G} \frac{1}{\binom{n-|F|}{n-|G|}}\right). \end{aligned}$$

It is easy to give a good lower estimate for the first factor. Let us investigate the second one:

$$1 - \sum_{G \in \mathcal{F}_i:\ F \subset G} \frac{1}{\binom{n-|F|}{n-|G|}}. \tag{20}$$

There is at most one $G \in \mathcal{F}_i$ of a given size by the condition of Theorem 2.1.6. Moreover $|F| < |G|$ holds. Therefore

$$1 - \frac{1}{\binom{n-|F|}{1}} - \frac{1}{\binom{n-|F|}{2}} - \frac{1}{\binom{n-|F|}{3}} - \ldots \tag{21}$$

is a lower estimate on (20). On the other hand, since the number of possible sizes (different from the size of F) is at most $2\alpha n$, the number of negative terms in (20) is at most $2\alpha n$.

Using this observation, a further lower estimate is obtained from (21):

$$1-\frac{1}{\binom{n-|F|}{1}}-\frac{1}{\binom{n-|F|}{2}}-\frac{2\alpha n}{\binom{n-|F|}{3}}.$$

Using the inequality $n-|F|\geq n\left(\frac{1}{2}-\alpha\right)$ in (21) gives the next estimate:

$$1-\frac{2}{n\left(1-2\alpha\right)}-\frac{2}{\left(n\left(\frac{1}{2}-\alpha\right)-1\right)^2}-\frac{12\alpha n}{\left(n\left(\frac{1}{2}-\alpha\right)-2\right)^3}. \tag{22}$$

Here $n\left(\frac{1}{2}-\alpha\right)-2\geq\frac{n}{4}$ holds by the conditions $0<\alpha\leq\frac{1}{8}, n\geq 16$. Substitute this into (22) and the final lower estimate yields that:

$$\begin{aligned} d(\mathcal{F}_i,F) &\geq 1-\frac{2}{n\left(1-2\alpha\right)}-\frac{2}{\left(\frac{n}{4}\right)^2}-\frac{12\alpha n}{\left(\frac{n}{4}\right)^3} \\ &= 1-\frac{2}{n\left(1-2\alpha\right)}-\frac{32}{n^2}-\frac{12\cdot 64\cdot\alpha}{n^2} \\ &\geq 1-\frac{2}{n\left(1-2\alpha\right)}-\frac{128}{n^2} \end{aligned}$$

is obtained where $\alpha\leq\frac{1}{8}$ was used. □

Lemma 2.4.2 *Under the assumptions of Lemma 2.4.1*

$$\frac{n!}{\binom{n}{\lfloor\frac{n}{2}\rfloor}}\left(1-\frac{2}{n(1-2\alpha)}-\frac{128}{n^2}\right)\leq\frac{c(\mathcal{F}_i)}{|\mathcal{F}_i|}$$

holds for every $\mathcal{F}_i\in F$.

Proof. We start with the inequality

$$\sum_{F\in\mathcal{F}_i}c(F)-\sum_{F,G\in\mathcal{F}_i,F\subset G}c(F,G)\leq c(\mathcal{F}_i). \tag{23}$$

This is an easy version of the sieve. Let us show it for the sake of completeness. If a chain $\mathcal{C}$ counted in $c(\mathcal{F}_i)$ satisfies $|\mathcal{C}\cap\mathcal{F}_i|=r>0$ then it is counted r times in the first term of the left hand side and $\binom{r}{2}$ times in the second term. $r-\binom{r}{2}\leq 1(0<r)$ shows that the left hand side of (23) counts every chain fewer times than the right hand side does.

Group together the terms on the left hand side of (23) containing F as the "smaller set".

$$\sum_{F\in\mathcal{F}_i}c(F)-\sum_{F,G\in\mathcal{F}_i,F\subset G}c(F,G)=\sum_{F\in\mathcal{F}_i}\left(c(F)-\sum_{G\in\mathcal{F}_i:\ F\subset G}c(F,G)\right) \tag{24}$$

$$=\sum_{F\in\mathcal{F}_i}d(\mathcal{F}_i,F) \tag{25}$$

(23) and (25) imply

$$\sum_{F\in\mathcal{F}_i} d(\mathcal{F}_i, F) \leq c(\mathcal{F}_i).$$

Using Lemma 2.4.1 for each $F \in \mathcal{F}_i$

$$|\mathcal{F}_i|\frac{n!}{\binom{n}{\lfloor\frac{n}{2}\rfloor}}\left(1-\frac{2}{n(1-2\alpha)}-\frac{128}{n^2}\right) \leq c(\mathcal{F}_i)$$

is obtained which is equivalent to the statement of the lemma. □

Lemma 2.4.2 shows that the function in (18) can be chosen to be

$$f(n,\alpha) = \frac{n!}{\binom{n}{\lfloor\frac{n}{2}\rfloor}}\left(1-\frac{2}{n(1-2\alpha)}-\frac{128}{n^2}\right).$$

Based on this, (19) gives

$$|\mathcal{F}| \leq \binom{n}{\lfloor\frac{n}{2}\rfloor}\frac{1}{1-\frac{2}{n(1-2\alpha)}-\frac{128}{n^2}}. \tag{26}$$

Applying the inequality

$$\frac{1}{1-x} \leq 1+x+2x^2 \quad \left(x \leq \frac{1}{2}\right)$$

in (26) with

$$x = \frac{2}{n(1-2\alpha)} + \frac{128}{n^2}$$

the upper estimate in Theorem 2.2.1 is obtained, finishing the proofs.

2.5 Lower Bound Construction

We include a proof of this known result because it plays a role in two different sections of this thesis. The proof, given in [21] is short and straightforward. This result provides the lower bounds in Theorems 2.1.3, 2.1.6, and 6.5.3. Recall that P_i is the i^{th} level of the Boolean lattice, where each element contains exactly i ones.

Lemma 2.5.1 *In $\mathfrak{B}_n$, there is a subset of $P_{\lfloor\frac{n}{2}\rfloor+1}$ with size at least $\frac{1}{n}\binom{n}{\lfloor\frac{n}{2}\rfloor}$ so this subset unioned with $P_{\lfloor\frac{n}{2}\rfloor}$ does not form any V_2.*

Proof. Given an element $\overrightarrow{x} \in \mathfrak{B}_n$, define $\eta(\overrightarrow{x}) = \sum_{x_i=1} i (\text{mod } n)$. Divide level $P_{\lfloor\frac{n}{2}\rfloor}$ into n equivalence classes as determined by η. Note that by the pigeonhole principle, at least one of these equivalence classes must have size at least $\frac{1}{n}|P_{\lfloor \frac{n}{2} \rfloor}|$. Let's say the equivalence class is S_ℓ; i.e. for $\overrightarrow{s} \in S_\ell$,

$$\eta(\overrightarrow{s}) = \sum_{s_i=1} i \pmod{n} = \ell.$$

Now we just need to show that $S_\ell \cup P_{\lfloor \frac{n}{2} \rfloor}$ forms no V_2. Let $\overrightarrow{r}$, $\overrightarrow{t}$ be sets in S_ℓ. We use a slight abuse of notation to regard sets in $\mathfrak{B}_n$ as vectors but still apply set operations, but we rely on the reader's familiarity with thinking about sets represented by their incidence vectors. Suppose that $\overrightarrow{r} \cap \overrightarrow{t} = \overrightarrow{y}$ for some $\overrightarrow{y} \in P_{\lfloor \frac{n}{2} \rfloor}$. Then there exists some distinct $m_1 \neq m_2 \in [n]$ so that $\overrightarrow{r} \backslash m_1 = \overrightarrow{y}$ and $\overrightarrow{t} \backslash m_2 = \overrightarrow{y}$. This implies that:

$$\sum_{y_j} j \pmod{n} = \Big(\sum_{r_i=1} i\Big) + m_1 \pmod{n} = \ell + m_1 \pmod{n}$$

and

$$\sum_{y_j} j \pmod{n} = \Big(\sum_{t_k=1} k\Big) + m_2 \pmod{n} = \ell + m_2 \pmod{n}$$

The second equalities in each of the lines above holds because $\overrightarrow{r}$ and $\overrightarrow{t} \in S_\ell$. However this creates a contradiction: $\ell + m_1 \,(\text{mod} n)$ cannot equal $\ell + m_2 \,(\text{mod} n)$ since m_1 and m_2 are distinct numbers between 0 and $n-1$. This shows that no such pair $\overrightarrow{r}$ and $\overrightarrow{t}$ with an intersection in $P_{\lfloor \frac{n}{2} \rfloor}$ can exist in S_ℓ, showing the desired result. $\square$

In Theorems 2.1.3 and 2.1.6 we find a large family which contains no V_2, by taking the middle layer of $\mathfrak{B}_n$ as well as the at least $\frac{1}{n}^{th}$ portion of the next layer up which belongs to S_ℓ. To apply this result to Theorem 6.5.3 we first have to note that by symmetry of the Boolean Lattice, that the lemma also shows there is a subset of the level below the middle with the appropriate size which when unioned with the middle layer creates no cherries. (See Section 6 for definitions).

2.6 Probabilistic Interpretation of the Proof

Many proofs have been given for Sperner's Theorem (Theorem 2.1.1). One of the most elegant is a probabilistic proof given in [2], p. 197. There they prove a lemma stating that for an antichain $\mathcal{A}$,

$$\sum_{A \in \mathcal{A}} \frac{1}{\binom{n}{|A|}} \leq 1.$$

Theorem 2.1.1 follows directly from this statement once the left hand side is interpreted as the expected number of times a uniformly random full chain passes through the antichain, and the denominator is maximized for $|A| = \binom{n}{\lfloor \frac{n}{2} \rfloor}$. Lemma 2.4.1 is our analogue of their lemma. However, we must compensate for the fact that a full chain may have unbounded intersection with a given component.

Let $\mathcal{F}$ be an induced V-free family and σ be a full chain chosen uniformly at random from the set of all full chains. After dividing both sides of (17) by $n!$ and recalling that σ can pass through at most one component of $\mathcal{F}$, we see that (17) can be seen as an application of the union bound:

$$\mathbb{P}(\sigma \text{ passes through } \mathcal{F}) = \sum_{i}^{K} \mathbb{P}(\sigma \text{ passes through } \mathcal{F}_i) = \sum_{i}^{K} \frac{c(\mathcal{F}_i)}{n!} \leq 1. \tag{27}$$

Section 2.4 is devoted to finding a value for the quantity $f(n, \alpha)$; Recall from (18) that $f(n, \alpha)$ satisfies:

$$f(n, \alpha) \leq \frac{c(\mathcal{F}_i)}{|\mathcal{F}_i|}.$$

Again rearranging terms, we can see that

$$\frac{|\mathcal{F}_i| f(n, \alpha)}{n!} \leq \frac{c(\mathcal{F}_i)}{n!} = \mathbb{P}(\sigma \text{ passes through } \mathcal{F}_i).$$

The proof of Theorem 2.1.6 now follows from substituting the value found in Section 2.4 for $f(n, \alpha)$ into (27).

CHAPTER III

COUNTING MATCHINGS AND INDEPENDENT SETS OF A FIXED SIZE

3.1 Introduction

Given a d-regular graph G on N vertices and a particular type of subgraph, a natural class of problems arises: "How many subgraphs of this type can G contain?" In this chapter we give upper bounds on the number of partial matchings of a fixed fractional size, and on the number of independent sets of a fixed size, in a general d-regular graph, and we show that our bounds are asymptotically matched at the logarithmic level by the graph consisting of $\frac{N}{2d}$ disjoint copies of $K_{d,d}$. (See [3] and [15] for graph theory basics.)

Let G be a bipartite graph on N vertices with partition classes A and B. Suppose that the degree sequence of one side is given by $\{r_i\}_{i=1}^{|A|}$. It follows from the well-known theorem of Brégman concerning the permanent of 0-1 matrices [6] (see also [2]) that we can bound the number of perfect matchings in G using the following expression:

Theorem 3.1.1 *(Brégman) Let $\mathcal{M}_{\text{perfect}}(G)$ be the set of perfect matchings in G. Then*

$$|\mathcal{M}_{\text{perfect}}(G)| \leq \prod_{i=1}^{|A|} (r_i!)^{\frac{1}{r_i}}.$$

When $r_i = d$ for all i and $|A|$ is divisible by d, equality in the above theorem is achieved by the graph consisting of $\frac{N}{2d}$ disjoint copies of the complete bipartite graph $K_{d,d}$, so we know that among d-regular bipartite graphs on N vertices, with $2d|N$, this graph contains the greatest number of perfect matchings.

Friedland et al. [19] propose an extension of this observation, which they call the "upper matching conjecture." They conjecture that among all d-regular bipartite graphs on N vertices, with $2d|N$, none has more matchings of size ℓ for $0 \leq \ell \leq \frac{N}{2}$ than the graph consisting of $\frac{N}{2d}$ copies of $K_{d,d}$. We provide asymptotic evidence for this conjecture, without assuming that G is bipartite. We will upper bound the logarithm of the number of

ℓ-matchings of a regular graph and show that, at the level of the leading term, this upper bound is achieved by the disjoint union of the appropriate number of copies of $K_{d,d}$.

Let $\mathcal{M}(G)$ be the set of all matchings of G. For $0 \leq \ell \leq \frac{N}{2}$, let $\mathcal{M}_\ell(G)$ be the set of matchings of size ℓ (i.e., with ℓ edges) and set $\mathcal{M}_\ell(G) = |\mathcal{M}_\ell(G)|$. Set $\alpha = \frac{2\ell}{N}$; with this definition for α, we can refer interchangeably to a matching of size ℓ or a matching whose size is an α-fraction of the maximum possible matching size. For ease of notation, we will refer to the graph consisting of $\frac{N}{2d}$ disjoint copies of $K_{d,d}$ as $DK_{N,d}$.

In what follows, $H(x) = -x\log x - (1-x)\log(1-x)$ is the usual binary entropy function. (All logarithms in this chapter are base 2.)

Theorem 3.1.2 *Let G be a d-regular graph on N vertices and ℓ an integer satisfying $0 \leq \ell \leq \frac{N}{2}$. Set $\alpha = \frac{2\ell}{N}$. The number of matchings in G of size ℓ satisfies*

$$\log(\mathcal{M}_\ell(G)) \leq \frac{N}{2}\left[\alpha \log d + H(\alpha)\right].$$

This bound is tight up to the first order term: for fixed $\alpha \in (0,1)$,

$$\log(\mathcal{M}_\ell(DK_{N,d})) \geq \frac{N}{2}\left[\alpha \log d + 2H(\alpha) + \alpha \log\left(\frac{\alpha}{e}\right) + \Omega\left(\frac{\log(d)}{d}\right)\right],$$

with the constant in the Ω term depending on α.

We show similar results for the number of independent sets in d-regular graphs. A point of departure for our consideration of independent sets is the following result of Kahn [25]. For any graph G write $\mathcal{I}(G)$ for the set of independent sets in G and write $i_t(G)$ for the set of independent sets of size t (i.e., with t vertices).

Theorem 3.1.3 *(Kahn) For any N-vertex, d-regular bipartite graph G,*

$$|\mathcal{I}(G)| \leq |\mathcal{I}(K_{d,d})|^{N/2d}.$$

Note that when $2d|N$, we have $|\mathcal{I}(K_{d,d})|^{N/2d} = |\mathcal{I}(DK_{N,d})|$. Kahn [25] proposes the following natural conjecture.

Conjecture 3.1.4 *For any N-vertex, d-regular graph G with $2d|N$ and any $0 \leq t \leq N/2$,*

$$i_t(G) \leq i_t(DK_{N,d}).$$

We provide asymptotic evidence for this conjecture.

Theorem 3.1.5 *For N-vertex, d-regular G, and $0 \le t \le N/2$,*

$$i_t(G) \le \begin{cases} 2^{\frac{N}{2}\left(H\left(\frac{2t}{N}\right)+\frac{2}{d}\right)} & \textit{in general} \\ 2^{\frac{N}{2}\left(H\left(\frac{2t}{N}\right)+\frac{1}{d}-\frac{\log e}{2d}\left(1-\frac{2t}{N}\right)^d\right)} & \textit{if } G \textit{ is bipartite} \\ 2^t\binom{\frac{N}{2}}{t} & \textit{if } G \textit{ has a perfect matching.} \end{cases} \tag{28}$$

On the other hand,

$$i_t(DK_{N,d}) \ge \begin{cases} \left(1-\frac{1}{c}\right)\binom{\frac{N}{2}}{t}2^{\frac{N}{2}\left(\frac{1}{d}-\frac{c}{d}\left(1-\frac{2t}{N}\right)^d\right)} & \textit{for any } c > 1 \\ \\ 2^t\binom{\frac{N}{2}}{t}\prod_{k=1}^{t-1}\left(1-\frac{2kd}{N}\right) & \textit{for } t \le \frac{N}{2d}. \end{cases} \tag{29}$$

In order to interpret the bounds in Theorem 3.1.5 we need to place them in context, comparing the upper and lower bounds in various regimes. To do this, we look at sequences of d-regular graphs on N vertices, allowing N, t and d to approach infinity at different rates relative to each other. First, similar to the result in Theorem 3.1.2, we bound the number of independent sets of a fixed fractional size. For general graphs, comparison with $DK_{N,d}$ gives agreement in the first order term at the logarithmic level, and agreement in the second order term when the graphs under consideration are bipartite.

Specifically, if N, d and t are sequences satisfying $t = \alpha\frac{N}{2}$ for some fixed $\alpha \in (0,1)$ and G is a sequence of N-vertex, d-regular graphs, then it is immediate from (28) that

$$\log i_t(G) \le \begin{cases} \frac{N}{2}\left[H\left(\alpha\right)+\frac{2}{d}\right] & \text{in general} \\ \\ \frac{N}{2}\left[H\left(\alpha\right)+\frac{1}{d}\right] & \text{if } G \text{ is bipartite,} \end{cases}$$

whereas if $N = \omega(d\log d)$ and $d = \omega(1)$ then taking $c = 2$ in the first bound of (29) and using Stirling's formula to analyze the behavior of $\binom{N/2}{\alpha N/2}$, we obtain the near matching lower bound

$$\begin{aligned} \log(i_t(DK_{N,d})) &\ge -1+\log\binom{\frac{N}{2}}{\alpha\frac{N}{2}}+\frac{N}{2}\left[\frac{1}{d}-\frac{2}{d}(1-\alpha)^d\right] \\ &= \frac{N}{2}\left[H(\alpha)+\frac{1}{d}-\frac{2}{d}(1-\alpha)^d\right]-\Omega(\log N) \\ &= \frac{N}{2}\left[H\left(\alpha\right)+\frac{1}{d}(1+o(1))\right]. \end{aligned}$$

We use $N = \omega(d \log d)$ and $d = \omega(1)$ to conclude $\log N = o(N/d)$ and $(1-\alpha)^d = o(1)$, respectively.

If $N = o\left(d/(1-\alpha)^d\right)$ and G is bipartite, then the gap between our bounds on $i_t(G)$ and $i_t(DK_{N,d})$ is even smaller, being only a multiplicative factor of $O(\sqrt{N})$; indeed, in this case (taking any $c = \omega(1)$) we obtain from the first bound of (29) that

$$i_t(DK_{N,d}) \geq (1 - o(1)) \binom{\frac{N}{2}}{t} 2^{\frac{N}{2}\left(H(\alpha)+\frac{1}{d}\right)}$$

So far we have considered sets of fixed fractional size. If we consider smaller sets, whose sizes scale with N/d rather than N, the final bounds in (28) and (29) come into play. Specifically, for any N, t and d

$$i_t(DK_{N,d}) \geq \begin{cases} \binom{\frac{N}{2}}{t} 2^{t(1+o(1))} & \text{if } t = o\left(\frac{N}{d}\right) \\ (1+o(1)) \binom{\frac{N}{2}}{t} 2^t & \text{if } t = o\left(\sqrt{\frac{N}{d}}\right) \end{cases} \tag{30}$$

Note that in the latter case, for G with a perfect matching we have $i_t(G) \leq (1+o(1)) i_t(DK_{N,d})$. To obtain (30) from (29) we use

$$\prod_{k=1}^{t-1} \left(1 - \frac{2kd}{N}\right) \geq e^{-\frac{4d}{N}\sum_{k=1}^{t-1} k} \geq e^{-\frac{2dt(t-1)}{N}}.$$

3.2 Counting Matchings

Given a graph G and a nonnegative real number λ, we can form weighted matchings of G by assigning each matching containing ℓ edges weight λ^ℓ. The weighted partition function, $Z^m_\lambda(G)$, gives the total weight of matchings. Formally,

$$Z^m_\lambda(G) := \sum_{m \in \mathcal{M}(G)} \lambda^{|m|} = \sum_{k=0}^{\frac{N}{2}} \mathcal{M}_k(G)\, \lambda^k.$$

We will prove Theorem 3.1.2 by showing a bound on the partition function, and then using that bound to limit the number of matchings of a particular weight (size).

Lemma 3.2.1 *For all d-regular graphs G, $Z^m_\lambda(G) \leq (1 + d\lambda)^{\frac{N}{2}}$*

This lemma is easily proven in the bipartite case; the difficulty arises when we want to prove the same bound for general graphs. Indeed, if G is a bipartite graph with bipartition classes A and B, we can easily see that the right hand side above counts a superset of weighted matchings. Elements in this superset are sets of edges no two of which are adjacent to the same element of A (but with no restriction on incidences with B).

Proof of Lemma 3.2.1 To prove this lemma, we will use the following result of Friedgut [18], which describes a weighted version of the information theoretic Shearer's Lemma. In the final sections of this chapter we give two alternate proofs for this lemma; one is a self contained combinatorial argument, and the other employs several analytic results.

Theorem 3.2.2 *(Friedgut) Let $H = (V, E)$ be a hypergraph, and $F_1, F_2, \ldots F_r$ subsets of V such that every $v \in V$ belongs to at least t of the sets F_i. Let H_i be the projection hypergraphs: $H_i = (V, E_i)$, where $E_i = \{e \cap F_i : e \in E\}$. For each edge $e \in E$, define $e_i = e \cap F_i$, and assign each e_i a nonnegative real weight $w_i(e_i)$. Then*

$$\Big(\sum_{e \in E} \prod_{i=1}^{r} w_i(e_i)\Big)^t \leq \prod_i \sum_{e_i \in E_i} w_i(e_i)^t$$

The first step in applying this theorem is to define appropriate variables. Let $G = (V, E)$ be a d-regular graph, with its vertex set $\{v_1, v_2, \ldots, v_N\}$. We will use G to form an associated matching hypergraph, $H = (E, \mathcal{M})$, where the vertex set of the hypergraph is the edge set of G, and $\mathcal{M}$ is the sets of matchings in G. Let F_i be the set of edges incident to a vertex $v_i \in V$. Note that each edge in E is covered twice by $\bigcup_{i=1}^{N} F_i$, so we may take $t = 2$. We define the trace sets, $E_i = \{F_i \cap m : m \in \mathcal{M}\}$, as the set of possible intersections of a matching with the set of edges incident with v_i. Let $m_i = m \cap F_i$. Then for all i, assign

$$w_i(m_i) = \begin{cases} 1 & \text{if } m_i = \emptyset \\ \sqrt{\lambda} & \text{else} \end{cases}$$

With these definitions we have $\sum_{m_i \in E_i} w_i(m_i)^2 = 1 + d\lambda$, and for a fixed m, $\prod\limits_i w_i(m_i) = \sqrt{\lambda}^{(2|m|)}$. Putting these expressions into Theorem 3.2.2, we have that

$$(Z_\lambda^m(G))^2 = \left(\sum_{m \in \mathcal{M}} \lambda^{|m|}\right)^2 \leq \prod_{i=1}^{N}(1 + d\lambda).$$

Therefore,

$$Z_\lambda^m(G) \leq (1 + d\lambda)^{\frac{N}{2}}.$$

$\square$

Proof of Theorem 3.1.2 We begin with the upper bound. We may assume $0 < \ell < N/2$, since the extreme cases $\ell = 0, N/2$ are obvious. For fixed ℓ, a single term of the partition function $Z_\lambda^m(G)$ is bounded by the whole sum, and so by Lemma 3.2.1 we have $\mathcal{M}_\ell(G)\lambda^\ell \leq Z_\lambda^m(G) \leq (1+d\lambda)^{\frac{N}{2}}$ and

$$\mathcal{M}_\ell(G) \leq (1+d\lambda)^{\frac{N}{2}} \left(\frac{1}{\lambda}\right)^\ell. \tag{31}$$

We take

$$\lambda = \frac{\ell}{d\left(\frac{N}{2} - \ell\right)}$$

to minimize the right hand side of (31) and obtain the upper bound in Theorem 3.1.2 (in the case $\ell = \frac{\alpha N}{2}$):

$$\begin{aligned} \log(\mathcal{M}_\ell(G)) &\leq \log \left(\frac{\frac{N}{2}}{\frac{N}{2} - \ell}\right)^{\frac{N}{2}} \left(\frac{d(\frac{N}{2} - \ell)}{\ell}\right)^\ell \\ &= \frac{N}{2}\left(\frac{2\ell}{N}\log d + H\left(2\ell/N\right)\right) \\ &= \frac{N}{2}\left(\alpha \log d + H(\alpha)\right). \end{aligned}$$

We now turn to the lower bound. We begin by observing

$$\mathcal{M}_\ell(DK_{N,d}) = \sum_{\substack{a_1,\dots a_{N/2d}: \\ 0 \leq a_i \leq d,\ \sum_i a_i = \ell}} \prod_{i=1}^{N/2d} \binom{d}{a_i}^2 a_i! \tag{32}$$

Here the a_i's are the sizes of the intersections of the matching with each of the components of $DK_{N,d}$, and the term $\binom{d}{a_i}^2 a_i!$ counts the number of matchings of size a_i in a single copy of $K_{d,d}$. (The binomial term represents the choice of a_i endvertices for the matching from each partition class, and the factorial term tells us how many ways there are to pair the endvertices from the top and bottom to form a matching.)

From Stirling's formula we know there is an absolute constant $c \geq 1$ such that for any $d \geq 1$ and $0 < a < d$,

$$\log\left(\binom{d}{a}^2 a!\right) \geq a \log d + a \log\frac{a}{d} - a\log e + 2H\left(\frac{a}{d}\right) d - \log cd, \tag{33}$$

and we may verify by hand that (33) holds also for $a = 0, d$. Combining (32) and (33) we see that $\log(\mathcal{M}_\ell(DK_{N,d}))$ is bounded below by

$$\frac{N}{2}\left(\frac{2\ell}{N}\log d - \frac{2\ell}{N}\log e - \frac{\log cd}{d} + \frac{2}{N}\sum_{i=1}^{N/2d}\left(a_i \log\left(\frac{a_i}{d}\right) + 2H\left(\frac{a_i}{d}\right)d\right)\right) \tag{34}$$

for any valid sequence of a_i's. To get our lower bound in the case $\ell = \alpha\frac{N}{2}$, we consider (34) a sequence of a_i's where each a_i is either $\lfloor \alpha d \rfloor$ or $\lceil \alpha d \rceil$. Note that by the mean value theorem, there is a constant $c_\alpha > 0$ such that both

$$\log\frac{\lceil \alpha d \rceil}{d},\ \log\frac{\lfloor \alpha d \rfloor}{d} \geq \log\alpha - \frac{c_\alpha}{d}$$

and

$$H\left(\frac{\lceil \alpha d \rceil}{d}\right),\ H\left(\frac{\lfloor \alpha d \rfloor}{d}\right) \geq H(\alpha) - \frac{c_\alpha}{d}.$$

(Here we use

$$\left|\frac{\lceil \alpha d \rceil}{d} - \alpha\right|,\ \left|\frac{\lfloor \alpha d \rfloor}{d} - \alpha\right| \leq \frac{1}{d}$$

and $\alpha \neq 0, 1$.) Putting these bounds into (34) we obtain

$$\log(\mathcal{M}_\ell(DK_{N,d})) \geq \frac{N}{2}\left(\alpha\log d + 2H(\alpha) + \alpha\log\left(\frac{\alpha}{e}\right) + \Omega\left(\frac{\log d}{d}\right)\right), \tag{35}$$

with the constant in the Ω term depending on α. □

3.3 A comment on the Lower Bound

We showed the lower bound from Theorem 3.1.2 by looking at a very specific way of forming a matching of size $\alpha\frac{N}{2}$, namely by giving every $K_{d,d}$ component as close to an α share of the available edges as possible. Another way to form a matching of size $k = \alpha\frac{N}{2}$ in $DK_{N,d}$ is to leave a certain number of the components empty, and distribute the remaining edges evenly in the other components. We will distribute βd edges, where $\alpha \leq \beta \leq 1$ is a fixed constant, to each of $\frac{\alpha N}{2\beta d}$ components, and then optimize with respect to β. The total number of matchings of this form gives us a lower bound for $M_{\frac{\alpha N}{2}}(DK_{N,d})$. Now we select which components are left empty, which gives us an additional term of the form $\binom{\frac{N}{2d}}{\frac{\alpha N}{2\beta d}}$. We make a tradeoff between making an additional choice of which components include edges and having a smaller number choices in each component. The purpose of this section is to show that this strategy is not asymptotically more successful than distributing approximately αd edges to each component. This also implies that a higher number of parameters assigned in this way, i.e. β edges to s components, and γ edges to t components with $\beta s + \gamma t = \alpha \frac{N}{2}$ will not give an improvement in the lower bound, as the outcome of the optimization of β and γ tells us that would be asymptotically equivalent to distribute all of the edges evenly in both of the subproblems.

We proceed in the proof, assuming that βd is an integer. Since this scheme will not yield an improvement when β is an integer there will also be no improvement when βd is taken to be an arbitrary real number.

Proposition 3.3.1 *Assuming that $\frac{\alpha}{\beta}$ is a fixed fraction in $(0,1)$,*

$$\begin{aligned}\log(M_{\frac{\alpha N}{2}}(DK_{N,d})) &\geq \frac{N}{2}\Big[\alpha\log(d) + \alpha\log\left(\frac{\beta}{e}\right) + \frac{2\alpha}{\beta d}H(\beta) + \frac{1}{d}H\left(\frac{\alpha}{\beta}\right) + O\left(\frac{\log(d)}{d}\right)\Big] \\ &\sim \frac{N}{2}\left(\alpha\log d + 2H(\alpha) + \alpha\log\left(\frac{\alpha}{e}\right) + \Omega\left(\frac{\log d}{d}\right)\right).\end{aligned}$$

Proof: There are

$$\binom{\frac{N}{2d}}{\frac{\alpha N}{2\beta d}}[\beta d!]^{\frac{\alpha N}{2\beta d}}\binom{d}{\beta d}^{2\frac{\alpha N}{2\beta d}}. \tag{36}$$

matchings with βd edges distributed to each of the appropriate number of components. The first binomial term represents the choice of affected components, whereas the second binomial term represents the choice of selected vertices on the top and bottom of these components. The factorial term tells us how many ways there are to pair the vertices from the top and bottom to form a matching. Taking the logarithm shows that:

$$\log(M_{\frac{\alpha N}{2}}(DK_{N,d})) \geq \frac{\alpha N}{2\beta d}\log(\beta d!) + \frac{\alpha N}{2\beta d}\log\left(\binom{d}{\beta d}\right) + \frac{N}{2d}\log\left(\binom{\frac{N}{2d}}{\frac{\alpha N}{2\beta d}}\right). \tag{37}$$

We use the fact that for any $0 < \gamma < 1$:

$$\binom{N}{\gamma N} \sim \frac{2^{H(\gamma)N}}{\sqrt{2\pi\gamma(1-\gamma)N}}. \tag{38}$$

Continuing from (37), we see that:

$$\begin{aligned}\log(M_{\frac{\alpha N}{2}}(DK_{N,d})) \geq \frac{N}{2}\Big[&\frac{\alpha}{\beta d}\log(\beta d!) + \frac{2\alpha}{\beta}H(\beta) + \frac{1}{d}H\left(\frac{\alpha}{\beta}\right) - \log\left(2\pi\beta(1-\beta)d\right)\Big] \\ &- \frac{1}{2}\log\left(2\pi\frac{\alpha}{\beta}\left(1-\frac{\alpha}{\beta}\right)\frac{N}{d}\right) \\ \geq \frac{N}{2}\Big[&\alpha\log(d) + \alpha\log\left(\frac{\beta}{e}\right) + \frac{\alpha}{\beta}H(\beta) + \frac{1}{d}H\left(\frac{\alpha}{\beta}\right) \\ &- \frac{\alpha}{2\beta d}\log(d) - \frac{\alpha}{2\beta d}\log(2\pi\beta) - \frac{\alpha}{\beta d}\log(1-\beta)\Big] \\ &- \frac{1}{2}\log\left(\frac{N}{2d}\right) - \frac{1}{2}\log\left(2\pi\frac{\alpha}{\beta}\left(1-\frac{\alpha}{\beta}\right)\right) \\ \geq \frac{N}{2}\Big[&\alpha\log(d) + \alpha\log\left(\frac{\beta}{e}\right) + \frac{\alpha}{\beta}H(\beta) + \frac{1}{d}H\left(\frac{\alpha}{\beta}\right) + O\left(\frac{\log(d)}{d}\right)\Big] \\ \geq \frac{N}{2}\Big[&\alpha\log\left(\frac{d}{e}\right) + \alpha\log(\beta) + \frac{\alpha}{\beta}H(\beta) + O\left(\frac{\log(d)}{d}\right)\Big].\end{aligned} \tag{39}$$

In the last line, it is easy to see that the $\frac{1}{d}H\left(\frac{\alpha}{\beta}\right)$ term can be grouped with the $O\left(\frac{\log(d)}{d}\right)$ term, as $H(x)$ is bounded by 1 for $0 < x < 1$. Now we can isolate the terms which depend on β and optimize.

Let $f(\beta) = \alpha\log(\beta) + \frac{\alpha}{\beta}H(\beta)$. Then:

$$\begin{aligned}f'(\beta) &= \frac{\alpha}{\beta}\left(1 + \log\left(\frac{1-\beta}{\beta}\right)\right) - \frac{\alpha H(\beta)}{\beta^2} \\ &= \frac{\alpha}{\beta}\left(1 + \log\left(\frac{1-\beta}{\beta}\right) - \frac{H(\beta)}{\beta}\right).\end{aligned} \tag{40}$$

We see that this expression is always negative. This implies that the maximum must occur when β is equal to one of its endpoints. However we have that $0 < \beta \leq \alpha$, so the optimum occurs when $\beta = \alpha$. Notice that when we substitute α for β in the final line of (39) we retrieve exactly the result from (35).

3.4 Proof of the Upper Bound in Theorem 3.1.2

In order to show the upper bound in Theorem 3.1.2, we will prove an upper bound which will hold for any d-regular graph G on N vertices. We want to use the statement from Lemma 3.2.1:

$$Z_\lambda^m(G) = \sum_{k=0}^{\frac{N}{2}} \mathcal{M}_k\lambda^k \leq (1+d\lambda)^{\frac{N}{2}}$$

to bound the number of matchings of size ℓ. We observe that or a fixed ℓ, a single term of the sum is bounded by the whole sum:

$$\mathcal{M}_\ell(G)\lambda^\ell \leq (1+d\lambda)^{\frac{N}{2}}.$$

Consequently,

$$\mathcal{M}_\ell(G) \leq (1+d\lambda)^{\frac{N}{2}}\left(\frac{1}{\lambda}\right)^\ell. \tag{41}$$

This inequality holds for all values of G, λ and ℓ, however we can optimize with respect to λ to make the inequality as tight as possible for a particular value of ℓ. This yields that

$$\lambda = \frac{\ell}{d\left(\frac{N}{2}-\ell\right)}. \tag{42}$$

Using $\ell = \alpha\frac{N}{2}$, and substituting this value for λ into 41, we get that

$$\begin{aligned}
\log(M_{\frac{\alpha N}{2}}(G)) &\leq \log\left(\left(1+\frac{\alpha}{1-\alpha}\right)^{\frac{N}{2}}\left(\frac{d-\alpha d}{\alpha}\right)^{\frac{\alpha N}{2}}\right) \\
&\leq \frac{N}{2}\left[\log\left(1+\frac{\alpha}{1-\alpha}\right)+\alpha\log(d)+\alpha\log\left(\frac{1-\alpha}{\alpha}\right)\right] \\
&\leq \frac{N}{2}[\alpha\log(d)+H(\alpha)]
\end{aligned} \tag{43}$$

3.5 Counting Independent Sets

In this section we prove the various assertions of Theorem 3.1.5. We begin with the second bound in (28). We use a result from [20], which states that for any $\lambda > 0$ and any d-regular N-vertex bipartite graph G, the weighted independent set partition function satisfies

$$Z_\lambda^{\text{ind}}(G) := \sum_{I\in\mathcal{I}(G)} \lambda^{|I|} \leq \left(2(1+\lambda)^d-1\right)^{\frac{N}{2d}}. \tag{44}$$

Choose λ so that $\frac{\lambda N}{2(1+\lambda)} = t$. Noting that $i_t(G)\lambda^{\frac{\lambda N}{2(1+\lambda)}}$ is the contribution to $Z_\lambda^{\text{ind}}(G)$ from independent sets of size t we have

$$\begin{aligned} i_t(G) &\leq \frac{Z_\lambda^{\text{ind}}(G)}{\lambda^{\frac{\lambda N}{2(1+\lambda)}}} \\ &\leq \frac{\left(2(1+\lambda)^d - 1\right)^{\frac{N}{2d}}}{\lambda^{\frac{\lambda N}{2(1+\lambda)}}} \\ &= 2^{\frac{N}{2d}} \left(\frac{1+\lambda}{\lambda^{\frac{\lambda}{1+\lambda}}}\right)^{N/2} \left(1 - \frac{1}{2(1+\lambda)^d - 1}\right)^{\frac{N}{2d}} \\ &= 2^{H\left(\frac{\lambda}{1+\lambda}\right)\frac{N}{2} + \frac{N}{2d}} e^{-\frac{N}{4d(1+\lambda)^d}} \\ &= 2^{H\left(\frac{2t}{N}\right)\frac{N}{2} + \frac{N}{2d} - \frac{N \log e}{4d}\left(1 - \frac{2t}{N}\right)^d}. \end{aligned} \tag{45}$$

We use (44) to make the critical substitution in (45).

To obtain the first bound in (28) we need the following analog of (44) for G not necessarily bipartite:

$$Z_\lambda^{\text{ind}}(G) \leq 2^{\frac{N}{d}} (1+\lambda)^{\frac{N}{2}}. \tag{46}$$

From (46) we easily obtain the claimed bound, following the steps of the derivation of the second bound in (28) from (44). We prove (46) by using a more general result on graph homomorphisms. For graphs $G = (V_1, E_1)$ and $H = (V_2, E_2)$ set

$$Hom(G,H) = \{f : V_1 \to V_2 \ : \ \{u,v\} \in E_1 \Rightarrow \{f(u), f(v)\} \in E_2\}.$$

That is, $Hom(G,H)$ is the set of graph homomorphisms from G to H. Fix a total order $\prec$ on $V(G)$. For each $v \in V(G)$, write $P_\prec(v)$ for $\{w \in V(G) : \{w,v\} \in E(G), w \prec v\}$ and $p_\prec(v)$ for $|P_\prec(v)|$. The following natural generalization of a theorem of Kahn is due to Galvin (see [33] for a proof).

Theorem 3.5.1 *For any d-regular and N-vertex graph G (not necessarily bipartite) and any total order $\prec$ on $V(G)$,*

$$|Hom(G,H)| \leq \prod_{v \in V(G)} |Hom(K_{p_\prec(v), p_\prec(v)}, H)|^{\frac{1}{d}}.$$

If G is bipartite with bipartition classes $\mathcal{E}$ and $\mathcal{O}$ and $\prec$ satisfies $u \prec v$ for all $u \in \mathcal{E}, v \in \mathcal{O}$ then Theorem 3.5.1 reduces to the main result of [20].

To prove (46), we first note that (by continuity) it is enough to prove the result for λ rational. Let C be an integer such that $C\lambda$ is also an integer, and let H_C be the graph which consists of an independent set of size $C\lambda$ and a complete looped graph on C vertices, with a complete bipartite graph joining the two. As described in [20] we have, for any graph G on N vertices,

$$|Hom(G, H_C)| = C^N Z_\lambda^{\mathrm{ind}}(G).$$

For G d-regular and N-vertex, we apply Theorem 3.5.1 twice to obtain

$$\begin{aligned}
Z_\lambda^{\mathrm{ind}}(G) &= \frac{|Hom(G, H_C)|}{C^N} \\
&\leq \frac{\prod_{v \in V(G)} |Hom(K_{p_\prec(v), p_\prec(v)}, H_C)|^{\frac{1}{d}}}{C^N} \\
&= \frac{\prod_{v \in V(G)} \left(C^{2p_\prec(v)} Z_\lambda^{\mathrm{ind}}(K_{p_\prec(v), p_\prec(v)})\right)^{\frac{1}{d}}}{C^N} \\
&\leq \frac{C^{\frac{2\sum_{v \in V(G)} p_\prec(v)}{d}} \prod_{v \in V(G)} \left(2(1+\lambda)^{p_\prec(v)}\right)^{\frac{1}{d}}}{C^N} \\
&= 2^{\frac{N}{d}} \frac{C^{\frac{2\sum_{v \in V(G)} p_\prec(v)}{d}} (1+\lambda)^{\frac{\sum_{v \in V(G)} p_\prec(v)}{d}}}{C^N}.
\end{aligned}$$

Now noting that

$$\sum_{v \in V(G)} p_\prec(v) = |E(G)| = \frac{Nd}{2}$$

we obtain

$$Z_\lambda(G) \leq 2^{\frac{N}{d}} (1+\lambda)^{\frac{N}{2}},$$

as claimed.

We now turn to the third bound in (28). Fix a perfect matching of G joining a set of vertices $A \subseteq V(G)$ of size $N/2$ to the set $B := V(G) \setminus A$. Let f be the bijection from subsets of A to subsets of B that moves the set along the chosen matching. Every independent set in G of size t is of the form $I_A \cup I_B$ where $I_A \subseteq A$, $I_B \subseteq B$, $f(A) \cap B = \emptyset$ and $|A| + |B| = t$. We therefore count all the independent sets of size t (and more) by choosing a subset of A of size t ($\binom{N/2}{t}$ choices) and a subset of this set to send to B via f (2^t choices).

To obtain the first bound in (29), we introduce a probabilistic framework and use Markov's inequality. If we divide a set of size $N/2$ into $N/2d$ blocks of size d and choose a uniform subset of size t, then the probability that this set misses a particular block is $\binom{N/2-d}{t}/\binom{N/2}{t}$. Let X be a random variable representing the number of blocks that the t-set misses. Let b_k equal the number of t-sets which miss exactly k blocks. Then $\mathbb{P}(X=k) = b_k/\binom{N/2}{t}$. Let χ_A be the indicator variable for the event A. Then

$$X = \sum_{i=0}^{\frac{N}{2d}} \chi_{\{\text{block i empty}\}}$$

and by linearity of expectation the expected number of blocks missed satisfies

$$\mu := \mathbb{E}(X) = \frac{N}{2d} \frac{\binom{\frac{N}{2}-d}{t}}{\binom{\frac{N}{2}}{t}} \leq \frac{N}{2d}\left(1 - \frac{2t}{N}\right)^d. \tag{47}$$

From Markov's inequality we have

$$\sum_{k=0}^{c\mu} \mathbb{P}(X=k) = \mathbb{P}(X \leq c\mu) \geq \left(1 - \frac{1}{c}\right).$$

We substitute the previously discussed value for $\mathbb{P}(X=k)$, yielding the inequality

$$\sum_{k=0}^{c\mu} b_k \geq \left(1 - \frac{1}{c}\right)\binom{\frac{N}{2}}{t}. \tag{48}$$

How many independent sets of size t does $DK_{N,d}$ have? To choose an independent set from $DK_{N,d}$ of size t, we first create a bipartition $\mathcal{E} \cup \mathcal{O}$ of $DK_{N,d}$ by choosing (arbitrarily) one of the bipartition classes of each of the $N/2d$ $K_{d,d}$'s of $DK_{N,d}$ to be in $\mathcal{E}$. We then choose a subset of $\mathcal{E}$ of size t. The number of subsets of $\mathcal{E}$ which have empty intersection with exactly k of the $K_{d,d}$'s that make up $DK_{N,d}$ is precisely b_k. Each of these subsets corresponds to $2^{\frac{N}{2d}-k}$ independent sets in $DK_{N,d}$. Combining this observation with (47) and (48) we obtain the first bound in (29):

$$\begin{aligned} i_t(DK_{N,d}) &= 2^{\frac{N}{2d}} \sum_{k \geq 0} 2^{-k} b_k \\ &\geq 2^{\frac{N}{2d} - c\mu} \sum_{k=0}^{c\mu} b_k \\ &\geq \left(1 - \frac{1}{c}\right)\binom{\frac{N}{2}}{t} 2^{\frac{N}{2}\left(\frac{1}{d} - \frac{c}{d}\left(1 - \frac{t}{M}\right)^d\right)}. \end{aligned}$$

Finally we turn to the second bound in (29). We obtain the claimed bound by considering all of the independent sets whose intersection with each component of $DK_{N,d}$ has size either

0 or 1:

$$i_t(DK_{N,d}) \geq (2d)^t \binom{\frac{N}{2d}}{t}.$$

After a little algebra, the right hand side above is seen to be exactly the right hand side of the second bound in (29).

3.6 Alternative Proof for Lemma 3.2.1

The purpose of this section is to provide a more self-contained proof of Lemma 3.2.1. Appealing to Theorem 3.2.2 uses a relatively strong theorem whose complex proof relies on counting mappings between multi-hypergraphs. The result and techniques of this section may be of independent combinatorial interest as well, as we originally pursued this result looking to improve a theorem by Ordentlich and Roth [35].

To proceed, we use a relevant hypergraph construction and introduce the notion of a hypergraph homomorphism.

Definition 3.6.1 *A hypergraph is **linear** if every for every pair of distinct edges E, F, $|E \cap F| \leq 1$.*

With a graph G (which may include loops) we associate a hypergraph $\mathscr{H}_G$ whose ground-set V is the set of edges of G and whose edge set is $\mathscr{E} = \{E_v : v \text{ a vertex of } G\}$ where E_v is the set of all edges of H that have v as an end vertex. The hypergraph $\mathscr{H}_G$ is always linear; if G is d-regular and has N vertices $\mathscr{H}_G$ is d-uniform and 2-regular with a ground set of size $\frac{Nd}{2}$.

Definition 3.6.2 *Let $\mathscr{H} = (V, \mathscr{E})$ be a hypergraph. A set $I \subseteq V$ is an **independent set** in $\mathscr{H}$ if for all $E \in \mathscr{E}, |I \cap E| \leq 1$.*

We write $\mathcal{I}(\mathscr{H})$ for the set of independent sets of $\mathscr{H}$, and for $\lambda \in \mathbb{C}$ set

$$Z_\lambda^{ind}(\mathscr{H}) = \sum_{I \in \mathcal{I}(\mathscr{H})} \lambda^{|I|}.$$

Note that there is a one-to-one correspondence between the independent sets of $\mathscr{H}_G$ and the matchings in G– in particular $Z_\lambda^m(G) = Z_\lambda^{ind}(\mathscr{H}_G)$. Theorem 3.2.1 will thus follow directly from the following theorem.

Theorem 3.6.3 *For a d-regular, t-uniform, linear hypergraph $\mathscr{H} = (V, \mathscr{E})$,*

$$Z_\lambda^{ind}(\mathscr{H}) \leq (1 + t\lambda)^{\frac{|V|}{t}}.$$

Definition 3.6.4 *For a fixed graph H (which may have loops) on vertex set $V(H)$ and a hypergraph $\mathscr{H}$ on vertex set V, say that a function $f : V \to V(H)$ is a **hypergraph homomorphism** from $\mathscr{H}$ to H if for all $x, y \in V$ with $x, y \in e$ for some edge e of $\mathscr{H}$, it holds that $f(x)f(y)$ is an edge in H.*

When the context is clear we will refer to hypergraph homomorphisms simply as homomorphisms. Let $\mathscr{H}$ be the set of a hypergraph; we define the notation:

$$\mathrm{Hom}\,(\mathscr{H}, H) = \{f : V(\mathscr{H}) \to V(H) : f \text{ is a homomorphism}\}.$$

In particular, $\mathrm{Hom}\,(\mathscr{H}, H_{ind}) = \mathcal{I}(\mathscr{H})$ where H_{ind} is the graph on two vertices with an edge between them and a loop at exactly one vertex. We will proceed by employing entropy and using a conditional version of Shearer's lemma that appears in [25]. For the background on the entropy function H as it applies to this problem see [20, 25].

Lemma 3.6.5 *Let $\mathbf{X} = (X_1, \ldots, X_n)$ be a random vector and let $\prec$ be a linear order on $\{1, \ldots, n\}$. Let $\mathcal{A}$ be a collection of subsets of $\{1, \ldots, n\}$ (possibly with repeats), such that for each $i \in \{1, \ldots, n\}$, i is contained in at least k members of $\mathcal{A}$. For each $A \in \mathcal{A}$ write $\mathbf{X}_A$ for $(X_i : i \in A)$ and write $i \prec A$ if $i \prec a$ for each $a \in A$. Then*

$$H(X) \leq \frac{1}{k} \sum_{A \in \mathcal{A}} H(X_A \mid \{X_i : i \prec A\}).$$

Let $\mathscr{H}$ and H be fixed. Let $\mathbf{f}$ be a uniform member of $\mathrm{Hom}\,(\mathscr{H}, H)$. We will use Lemma 3.6.5 to bound the entropy of $\mathbf{f}$ and therefore $|\mathrm{Hom}\,(\mathscr{H}, H)|$. The main issue is to find an appropriate covering collection $\mathcal{A}$. We follow a recent idea of Kahn [25]. Let $\prec$ be a linear order on V. For each edge e in $\mathscr{H}$ denote by $\mathrm{top}\,(e)$ the maximal element in e with respect to $\prec$ and for each $x \in V$ write $m(x)$ for the number of edges e for which $x = \mathrm{top}\,(e)$ (so $0 \leq m(x) \leq d$). Set

$$P_\prec(x) = \{x' \in V \setminus \{x\} : x' \in e \text{ for some edge } e \text{ with } \mathrm{top}\,(e) = x\}.$$

We are now ready to define our covering family $\mathcal{A}$. For each $x \in V$ take one copy of $P_{\prec}(x)$ and $d - m(x)$ copies of $\{x\}$. We claim that each $x \in V$ appears in exactly d members of $\mathcal{A}$. Indeed, x appears once for each edge e with $x \in e$ and $\mathrm{top}\,(x) \neq e$, giving a contribution of $d - m(x)$, and it appears $m(x)$ times as a singleton.

For notational convenience we suppress dependence on $\prec$ in what follows. For each $x \in V$ and each possible restriction R_x of $\mathbf{f}$ to $P(x)$, set

$$q_x(R_x) = \mathbf{P}(\mathbf{f}|_{P(x)} = R_x)$$

and let r_x^{ext} be the number of extensions of the partial homomorphism R_x on $P(x)$ to a partial homomorphism on $P(x) \cup \{x\}$.

Applying Lemma 3.6.5 to obtain the first inequality, and with the rest of the steps justified below, we obtain

$$\begin{aligned}
H(\mathbf{f}) &\leq \frac{1}{d}\sum_{x \in V}(H(\mathbf{f}|_{P(x)}|\{y : y \prec P(x)\}) + m(x)H(\mathbf{f}|_{\{x\}}|\{\mathbf{f}(y) : y \prec x\})) \\
&\leq \frac{1}{d}\sum_{x \in V}(H(\mathbf{f}|_{P(x)}) + m(x)H(\mathbf{f}|_{\{x\}}|\mathbf{f}|_{P(x)})) &(49)\\
&= \frac{1}{d}\sum_{x \in V}\sum_{R_x}\left(\begin{array}{l} q_x(R_x)\log\frac{1}{q_x(R_x)} + \\ q_x(R_x)m(x)H(\mathbf{f}|_{\{x\}}|\mathbf{f}|_{P(x)} = R_x) \end{array}\right) &(50)\\
&\leq \frac{1}{d}\sum_{x \in V}\sum_{R_x}\left(q_x(R_x)\log\frac{1}{q_x(R_x)} + q_x(R_x)\log(r_x^{ext})^{m(x)}\right) &(51)\\
&= \frac{1}{d}\sum_{x \in V}\sum_{R_x} q_x(R_x)\log\frac{(r_x^{ext})^{m(x)}}{q_x(R_x)} &(52)\\
&\leq \frac{1}{d}\sum_{x \in V}\log\sum_{R_x}(r_x^{ext})^{m(x)}. &(53)
\end{aligned}$$

In (49) we have relaxed the condition in the entropy expressions somewhat. In (50) we use the definition of the entropy function to expand out the terms in the sum. In (51) we use the fact that $H(\mathbf{X}) \leq \log|range(\mathbf{X})|$ for any random variable $\mathbf{X}$ taking on finitely many values. In (52) we gather together the terms inside the sum. In (53) we apply Jensen's inequality.

Since $H(\mathbf{f}) = |\log|\mathrm{Hom}\,(\mathscr{H}, H)||$, we have established that for any graph H and any

t-uniform, d-regular linear hypergraph $\mathscr{H}$,

$$\log |\mathrm{Hom}\,(\mathscr{H}, H)| \leq \frac{1}{d} \sum_{x \in V} \log \sum_{R_x} (r_x^{ext})^{m(x)}. \tag{54}$$

In general, this is not a nice expression to analyze. Fortunately, the graphs H that we will be looking at have a nice structure. Specifically, for $\lambda > 0$ and integer C with $C\lambda \in \mathbb{N}$ (by continuity we may assume that λ is rational) define a graph $H(C, \lambda)$ on a vertex set of size $C(1 + \lambda)$ by partitioning the vertex set into a set V_{in} of size $C\lambda$ and a set V_{out} of size C, putting in all edges between V_{in} and V_{out}, all edges between pairs of vertices in V_{out} (including all possible loops) and no edges between pairs of vertices in V_{in}. (A similar idea was employed in [20].) We analyze the quantity $\sum_{R_x} (r_x^{ext})^{m(x)}$ for $H(C, \lambda)$. There are $C^{(t-1)m(x)}$ restrictions to $P(x)$ which place no constraint on the value at x (these are the restrictions which only use vertices from V_{out}), so from these restrictions we get a contribution of $(C(1 + \lambda))^{m(x)}$. The total number of restrictions to $P(x)$ is $(C^{t-1} + (t - 1)C\lambda C^{t-2})^{m(x)}$ (for each e with $x = \mathrm{top}\,(e)$, there are C^{t-1} restrictions with images drawn solely from V_{out}; there are $t-1$ choices for a vertex to get a value from V_{in}, $C\lambda$ choices for the value, and C choices for the remaining $t - 2$ vertices; and there are $m(x)$ such edges, which may be treated independently). Therefore there are $(C^{t-1}+(t-1)C\lambda C^{t-2})^{m(x)} - C^{(t-1)m(x)}$ restriction which place a constraint on x (the constraint being that x may take on no value from V_{in}), so each of these contributes $C^{m(x)}$ to the count. It follows that the total sum is

$$C^{(t-1)m(x)}(C(1 + \lambda))^{m(x)} + ((C^{t-1} + (t - 1)C\lambda C^{t-2})^{m(x)} - C^{(t-1)m(x)})C^{m(x)}$$

which simplifies to

$$C^{tm(x)}((1 + \lambda)^{m(x)} + (1 + (t - 1)\lambda)^{m(x)} - 1).$$

For $t \geq 2$ this is at most

$$C^{tm(x)}(1 + \lambda + (t - 1)\lambda)^{m(x)} = C^{tm(x)}(1 + t\lambda)^{m(x)}.$$

Inserting into (53) we obtain

$$\begin{aligned}
|\mathrm{Hom}\,(\mathscr{H}, H(C,\lambda))| &\leq \exp\left\{\frac{1}{d}\sum_{x\in V}\log C^{tm(x)}(1+t\lambda)^{m(x)}\right\} \\
&= \prod_{x\in V} C^{m(x)}(1+t\lambda)^{\frac{m(x)}{t}} \\
&= C(1+t\lambda)^{\sum_{x\in V} m(x)} \\
&= C^N(1+t\lambda)^{\frac{N}{t}} \qquad (55)
\end{aligned}$$

where in (55) we use that $\sum_{x\in V(\mathcal{H})} m(x) = Nd/t$ (each edge has a unique top element and so is counted exactly once in the sum of the $m(x)$'s).

We observe now that

$$|\mathrm{Hom}\,(\mathscr{H}, H(C,\lambda))| = C^N Z_\lambda^{ind}(\mathscr{H}). \qquad (56)$$

Indeed, we can partition $\mathrm{Hom}\,(\mathscr{H}, H(C,\lambda))$ into classes indexed by $\mathcal{I}(\mathscr{H})$ by noting that the image of V_{in} in the homomorphism is an independent set in $\mathscr{H}$. The number of homomorphisms that go into the class of a particular $I \in \mathcal{I}(\mathscr{H})$ is $C^N\lambda^{|I|}$, verifying (56). Combining (55) and (56) we obtain

$$Z_\lambda^{ind}(\mathscr{H}) \leq (1+t\lambda)^{\frac{N}{t}},$$

as claimed.

3.7 An Alternative approach to Lemma 3.2.1

Another proof for Lemma 3.2.1 was suggested to us by Leonid Gurvits. It has the advantages of being short and admitting a slightly more general result, but relies on some deep results of Heilmann and Lieb [23].

Lemma 3.7.1 *Let $\overline{d}$ be the average degree of a general graph G on N vertices. Then:*

$$Z_\lambda^m(G) \leq (1+\overline{d}\lambda)^{\frac{N}{2}}$$

Proof. Let $P(x)$ be a polynomial of degree D. If P has nonnegative coefficients and real roots then it is well known that $P(x)^{\frac{1}{D}}$ is concave in x. (for instance, see [4] p. 454). Using the concavity of $P(x)$ we can say that:

$$\frac{P(x)^{\frac{1}{D}} - P(0)^{\frac{1}{D}}}{x} \leq \frac{1}{D} P'(x).$$

Rearranging terms we see that:

$$P(x) \leq \left[P(0)^{\frac{1}{D}} + \frac{1}{D} P'(0) x \right]^D . \tag{57}$$

We will be applying this fact with the degree $\frac{N}{2}$ polynomial

$$P(\lambda) = Z_\lambda^m(G) = \sum_{k=0}^{\frac{N}{2}} \mathcal{M}_k(G)\, \lambda^k.$$

Since the coefficients in the matching polynomial count the number of matchings of a particular size, it is clear that they are always nonnegative. Heilmann and Lieb [23] have shown that $Z_\lambda^m(G)$ has real roots for every G.

$Z(0) = 1$, since the constant term is the number of matchings of size zero. $Z'(0)$ gives the number of matchings of size one, but that is precisely the number of edges. We can write the number of edges as $\frac{N\bar{d}}{2}$. Then we have:

$$\left[1^{\frac{2}{N}} + \frac{2}{N} \left(\frac{N\bar{d}}{2} \lambda \right) \right]^{\frac{N}{2}}$$

If G is regular than when $\frac{N\bar{d}}{2}$ is plugged into (57), we will reclaim the result of Lemma 3.2.1.

CHAPTER IV

COUNTING LINEAR EXTENSIONS OF POSETS

4.1 History

Definition 4.1.1 *Given a poset $(P, \preceq)$, a* ***linear extension*** *of P is a total ordering $(T, \leq)$ on the ground set of P which preserves the relationship $\preceq$. In other words if $x \preceq y$ in P, then $x \leq y$ in T.*

Let $L(P)$ denote the number of linear extensions of a poset P. In this section, we will give a result counting the number of linear extensions of bipartite posets with uniformly bounded up and down degrees. The final section of this chapter introduces $F_{n,k}$, the poset of partially defined functions, and gives asymptotic estimations for $\frac{\log(L(F_{n,k}))}{|F_{n,k}|}$.

This quantity is useful in several applications. For instance, choosing a linear extension uniformly at random from the set of all linear extensions of a given poset is a classic problem in uniform sampling and approximate counting. To this end, several groups have developed Markov chains which act on the space of linear extensions (see [27] and [9]). Here we are interested in asymptotic enumeration.

Proposition 4.1.2 *Let P be a ranked poset with height K. If we enumerate the levels of P as P_i for $i = 1, \ldots, K$, with associated rank sequence $\{r_i\}_{i=1}^K$, a lower bound for $L(P)$ is given by the expression*

$$\prod_{i=1}^{K} r_i! \leq L(P) \tag{58}$$

Proof. The right hand side counts the number of linear extensions which can be formed by ordering each level set individually and then concatenating the orderings of each level putting all elements of a given level before all of the elements in a higher level.

The first non-trivial upper bound for $L(P)$ for ranked posets was given by Sha and Kleitman [40]. They proved that:

Theorem 4.1.3 *(Sha and Kleitman) For a ranked poset, P, as above:*

$$L(P) \leq \prod_{i=1}^{K} (r_i)^{r_i}.$$

Definition 4.1.4 *A **regular** poset is a graded poset where the up and down degrees of a vertex are completely determined by its rank.*

Stanley evidently raised the problem of how many linear extensions of the Boolean Lattice there are (see [40]), although the question has been raised independently in other contexts. This problem has generated much of the interest in this area. While they were calculating asymptotics for $L(\mathfrak{B}_n)$, Brightwell and Tetali [8] gave a stronger upper bound for the number of linear extensions in the case where the poset P is regular. Let u_j be the up degree for all vertices on level j, and similarly let d_j be the down degree of all vertices on level j. Let r be the harmonic average of the down degrees:

$$r = \sum_{j=2}^{K} \frac{r_{j-1}}{d_j}. \tag{59}$$

Theorem 4.1.5 *(Brightwell and Tetali) For a regular poset, P with $|P| = N$, as above:*

$$L(P) \leq \left(\prod_{i=1}^{K} r_i! \left(\frac{2e(K-1)N}{r} \right)^r \right)$$

A key ingredient in the proof of Theorem 4.1.5 is a two level theorem involving order preserving maps (see Theorem 4.1.8 below), a structure which generalizes linear extensions. In the remainder of this section, we prove an extension of this two level result.

Definition 4.1.6 *For a poset $P(X, \preceq)$ and an $M \in N$, an **order-preserving map**, $\hat{f}$, is a mapping from P to $[0, \ldots, M]$, which respects the ordering of P; i.e. if $x \preceq y$ in P, then $\hat{f}(x) \leq \hat{f}(y)$. We can also think of order-preserving maps as a poset homomorphism from P to C_M, the chain on M vertices.*

Let $O_M(P)$ be the number of order preserving maps from P to $[0, \ldots, M]$. The quantity $O_M(P)$ is relevant for asyptomptically determining the number of linear extensions, as was first proven by Shepp [43, 44], where he showed the following result:

Lemma 4.1.7 *For any N-element poset P,*

$$\lim_{M->\infty} \frac{O_M(P)}{M^N} = \frac{L(P)}{N!}.$$

Let P be an N-element ranked regular poset with two levels, A and B. Let $|A| = a$, and $|B| = b$. Let the up and down degrees be defined respectively as $u = d_{up}(y)$ for all $y \in B$, and $d = d_{down}(x)$ for all $x \in A$. In the course of their proof of Theorem 4.1.5, Brightwell and Tetali [8] show that:

Theorem 4.1.8 *(Brightwell, Tetali)*

$$O_M(P) \leq \left(\sum_{j=1}^{M} \Big[[M-j+1]^u - [M-j]^u \Big] j^d \right)^{\frac{N}{d+u}} = O_M(K_{d,u})^{\frac{N}{d+u}}.$$

We show a generalization of this result, with a weaker degree hypothesis:

Theorem 4.1.9 *Given a poset P on two levels A and B, so that each element of A is below at most U elements of B, each element of B is above at least D elements of A.*

$$O_M(P) \leq \left(\sum_{j=1}^{M} \left[[M-j+1]^U - [M-j]^U \right] j^D \right)^{\frac{a}{D}}. \tag{60}$$

Proof. (following the proof of Brightwell and Tetali [8]) Let f be a uniform random variable chosen over all order-preserving maps between P and $[0, \ldots, M]$. Then we will have that $\log(O_M(P)) = H(f)$. We will give an upper bound for the entropy of f, thereby establishing the desired bound on number of order-preserving maps.

First we must establish some notation:

- For $X \subseteq V(P)$ when we restrict f to X we will use the notation $f_X = f|_X$.
- For $x \in A$, let $N_{up}(x) = \{y \in B | x \preceq y\}$, (note that since P is bipartite, this is merely the neighborhood of x. We use this extra notation as a reminder of the orientation of the poset).
- For $x \in A$, let $u_x = |N_{up}(x)|$ and recall that $u_x \leq U$ for all x.

- Let $Y_x = \min(f_{N_{up}(x)})$. This is the lowest position of any of x's neighbors in the order-preserving map determined by f.

- For $x \in A$ and $j \in [m]$, let $\mathbb{P}(Y_x = j)$ be denoted by $p_x(j)$. This is the probability that x's lowest neighbor is mapped to the value j.

Using this notation and the chain rule for entropy, we see that:

$$\log(O_M(P)) = H(f) = H(f_B) + H(f_A|f_B). \tag{61}$$

We will deal with each of these terms separately; first getting a bound on $H(f_B)$. Let Λ be the collection of sets of the form $\{i \,|\, y_i \in N_{up}(x)\}_{x \in A}$. Note that Λ covers every element in A at least D times by the degree requirements. Therefore, this set system satisfies the hypothesis of Shearer's Lemma. This allows us to write the term $H(f_B)$ as a sum in terms of elements of A. More precisely:

$$H(f_B) \leq \frac{1}{D} \sum_{\lambda \in \Lambda} H(f_\lambda) = \frac{1}{D} \sum_{x \in A} H(f_{N_{up}(x)}). \tag{62}$$

Now we would like an estimate on $H(f_{N_{up}(x)})$ which is uniform over all $x \in A$. First we note that $f_{N_{up}}(x)$ tells us the position of all of x's neighbors, so it certainly determines Y_x, the height of the lowest neighbor. Therefore, when we condition on U_x, we have that $H(Y_x|f_{N_{up}(x)}) = 0$. This allows for a simplified expression when we apply the chain rule of entropy, giving that:

$$\begin{aligned} H(f_{N_{up}(x)}) &= H(Y_x) + H(f_{N_{up}(x)}|Y_x) \\ &= \sum_{j=1}^{M} p_x(j) \log\left(\frac{1}{p_x(j)}\right) + \sum_{j=1}^{M} p_x(j) H(f_{N_{up}(x)}|Y_x = j). \end{aligned} \tag{63}$$

There are $(M-j+1)^{u_x} - (M-j)^{u_x}$ values that $f_{N_{up}(x)}$ can take which are consistent with $\min(f_{N_{up}(x)} = Y_x = j)$. If the lowest neighbor is in position j,there are $(M-j+1)^{u_x}$ ways to place $N_{up}(x)$ above j, and we subtract off the $(M-j)^{u_x}$ of them which do not use the j^{th} position. Since each u_x is uniformly bounded by U and $(M-j+1)^\ell - (M-j)^\ell$ is an increasing function in ℓ, we have that

$$\begin{aligned}\text{Range}(H(f_{N_{up}(x)}|Y_x = j)) &= (M-j+1)^{u_x} - (M-j)^{u_x} \\ &\le (M-j+1)^U - (M-j)^U.\end{aligned} \tag{64}$$

Since $H(X) \le \log(|\text{Range}(X)|)$, this gives us a bound on $H(f_{N_{up}(x)}|Y_x = j)$. This implies that

$$H(f_B) \le \frac{1}{D}\sum_{j=1}^{M} p_x(j)\log\left(\frac{1}{p_x(j)}\right) + p_x(j)(M-j+1)^U - (M-j)^U. \tag{65}$$

We return to $H(f_A)$, the remaining term in 61. Enumerating the set A as $(x_1, x_2, \ldots, x_a)$, we can write $f_A = (f(x_1), f(x_2), \ldots, f(x_a))$. This notation allows us to use the conditional subadditivity of entropy to say that:

$$\begin{aligned}H(f_A|f_B) &\le \sum_{x\in A} H(f(x)|f_B) \\ &\le \sum_{x\in A} H(f(X)|Y_x) \\ &= \sum_{x\in A}\sum_{j=1}^{M} p_x(j)H(f(x)|Y_x = j) \\ &\le \sum_{x\in A}\sum_{j=1}^{M} p_x(j)\log(j).\end{aligned} \tag{66}$$

The second inequality holds because f_B determines Y_x and the final inequality uses the fact that $\text{Range}((f(x)|y_x = j)) = j$.

Returning to (61), and combining the estimates for $H(f_B)$ and $H(f_A|f_B)$ from (65) and (66) respectively, we have that:

$$\begin{aligned}&\log(O_M(P)) = H(f_B) + H(f_A|f_B) \\ &\le \frac{1}{D}\left[\sum_{x\in A}\Big[\sum_{j=1}^{M} p_x(j)\log\left(\frac{1}{p_x(j)}\right) + p_x(j)\log\big((M-j+1)^U - (M-j)^U\big) + p_x(j)D\log(j)\Big]\right].\end{aligned} \tag{67}$$

Since $\log(x)$ is concave, combining terms into one logarithmic factor allows us to directly apply Jensen's inequality to the sum.

$$\begin{aligned}
\log(O_M(P)) &\leq \frac{1}{D}\left[\sum_{x\in A}\Big[\sum_{j=1}^{M} p_x(j)\log\Big(\frac{(M-j+1)^U-(M-j)^U)\,j^D}{p_x(j)}\Big)\Big]\right] \\
&\leq \frac{1}{D}\left[\sum_{x\in A}\Big[\sum_{j=1}^{M}\log\left([(M-j+1)^U-(M-j)^U]\,j^D\right)\Big]\right] \qquad (68) \\
&\leq \frac{a}{D}\left[\sum_{j=0}^{M}\log\left([(M-j+1)^U-(M-j)^U]\,j^D\right)\right].
\end{aligned}$$

Exponentiating both sides gives us the desired result. □

A drawback of this result is that it is only valid for ranked posets which can be written as the union of two levels with uniform degree bounds. If we have a bipartite ranked poset with different up and down degrees on each level, this technique does not allow us to use the additional information from those degree bounds, and in fact may give a very weak result if the degrees have wide variation. It is still unclear whether this information can be used in an entropy based approach, or otherwise to give a tighter result.

4.2 Linear Extensions of the Function Poset

For $n, k \in \mathbb{N}$ with $k \geq 1$, let $F_{n,k}$ represent the poset of partially defined functions from $[n]$ to $[k]$ ordered by inclusion. We will use $\mathcal{D}(f)$ to denote the domain of a function f. Functions f and g satisfy that $f \preceq g$ if and only if $\mathcal{D}(f) \subseteq \mathcal{D}(g)$ and $f(x) = g(x)$, for all $x \in \mathcal{D}(f)$. In other words, $f \preceq g$ if and only if g is an extension of f. $F_{n,k}$ is a ranked poset on $n+1$ levels, with $|F_{n,k}| = (k+1)^n$.

There are two primary special cases of the function poset. For all values of n, $F_{n,1}$ is isomorphic to the Boolean lattice B_n. We see that each element in $F_{n,1}$ can be thought of a string using the alphabet $\{0, 1\}$ where 0 represents the undefined character. In order to introduce the second special case, recall the definition of dual posets and Hasse diagrams.

The dual of a ranked poset, $P(X, \preceq)$, is a poset, P^* on the same ground set with all of the relations reversed. The Hasse diagram of a poset is a reduced schematic diagram showing edges between vertices only if $x \lessdot y$. Taking the dual can be thought of visually

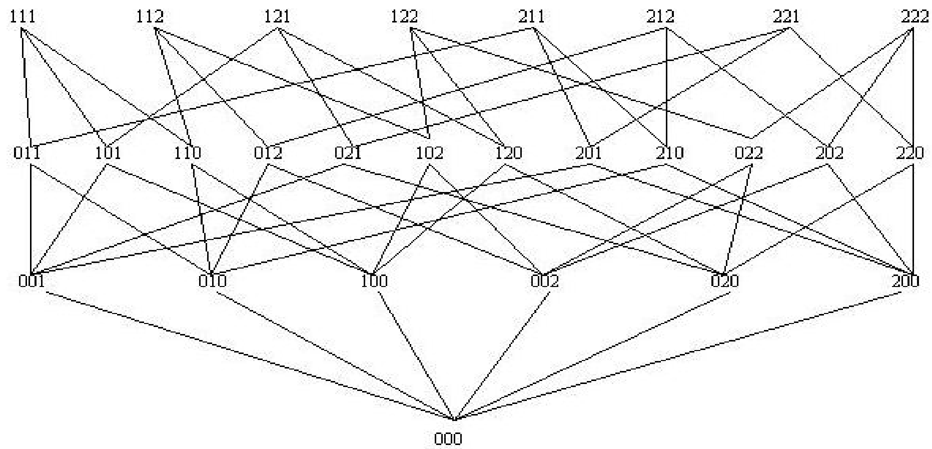

Figure 5: The Hasse diagram of $F_{3,2}$. 0 represents the undefined character.

as turning the poset upside down and reversing all of the edges in the Hasse diagram. (See [16] for more information on duals and poset definitions). Many properties of a poset are preserved under taking the dual, notably the number of linear extensions and the number of antichains. Suppose that $P(X, \preceq)$ is a poset with rank function $\{r_P\}$ which takes values from $i = 1, \ldots, K$. The rank function of the dual is $r(p) := K - r_P(p)$ for $p \in X$.

$F_{n,2}$ is isomorphic to the dual of the cubic poset, Q_n, which is the poset of lower dimensional faces of B_n (not including $\emptyset$) ordered by inclusion. We can think of Q_n as the set of n-tuples taking values in 0,1,2 ordered by $\vec{b} \leq \vec{c}$ if and only if $c_i = 2$ or $b_i = c_i$ for all $i \in [n]$. Given a set $I \in [n]$, for $i \in I$ we fix values $a_i = \alpha_i \in \{0, 1\}$. If we then let a_i range over 0 and 1 for all $i \notin I$, this determines a face of B_n.

The $n + 1$ levels of $F_{n,k}$ are indexed by the values $i \in \{0, \ldots, n\}$. The set of functions with exactly i values defined forms the i^{th} level, A_i. The size of the A_i is $n_i = k^i \binom{n}{i}$. The levels are indexed so that levels closer to the "bottom" of the poset have lower indices. It is easy to check that the rank sequence $\{r_i\}_{i=0}^{n}$ is unimodal, with the mode occurring at rank $\left\lfloor \frac{k}{k+1} \right\rfloor n$.

$F_{n,k}$ is a graded poset: there is a unique minimum element, and all maximal elements have the same rank. Additionally, the up and down degrees of an element are determined by the level it belongs to. If f is a member of level i it is defined in exactly i coordinates, with the remaining $n-i$ coordinates left undefined. The down degree d_i of level i is equal to i for each level. To select a downward neighbor of f, we just need to pick any of the i coordinates where it is undefined and replace the value there with the undefined character. The up degree of level i for $i = 0, \ldots, n-1$ is $u_i = k(n-i)$: to find an upward neighbor of f, we first choose one of the $n-i$ coordinates which is undefined in f and assign it one of the k allowed values.

Recall that for $F_{n,k}$, the number of levels in the poset is $K = n+1$, the rank function is $r_i = \binom{n}{i}k^i$, and $|F_{n,k}| = (k+1)^n$. For ease of notation let

$$\mathcal{L} = \frac{1}{(k+1)^n} \log \left(\prod_{i=0}^{n} \left(\binom{n}{i} k^i \right)! \right).$$

Theorem 4.2.1

$$\mathcal{L} \leq \frac{\log(L(F_{n,k}))}{|F_{n,k}|} \leq \mathcal{L} + O\left(\frac{\log n}{n}\right).$$

More specifically,

$$n\log(k+1) - \log(n+1) \leq \frac{\log(L(F_{n,k}))}{|F_{n,k}|} \leq n\log(k+1) - \log(e) + o(1).$$

Proof. Using the results from (58) and Theorem 4.1.5, for a ranked poset P with levels indexed by $1, \ldots, K$ and rank sequence $\{r_i\}_{i=1}^K$ we know that:

$$\frac{1}{|P|} \log \left(\prod_{i=1}^{K} r_i! \right) \leq \frac{1}{|P|} \log\left(L(P)\right) \leq \frac{1}{|P|} \log \left(\left(\prod_{i=1}^{K} r_i! \left(\frac{2e(K-1)N}{r} \right)^r \right) \right). \tag{69}$$

(Note that even though these bounds are originally given for posets with K levels with the indexing beginning with one, we can reindex to start with 0. In $F_{n,k}$, there is a unique element of rank 0, so $r_0 = 1$. This gives us a multiplicative factor of 1 in both the upper and lower bounds, so it is equivalent for us to begin our indices with either $i = 1$ or $i = 0$).

We proceed with this proof by giving an expansion for $\mathcal{L}$ since this term appears in both the upper and lower bounds. Then we show that the error term in the upper bound, $\frac{\log\left(\left(\frac{2e(K-1)N}{r}\right)^r\right)}{|F_{n,k}|}$ has a smaller order; in fact we will show that it has order $O\left(\frac{\log n}{n}\right)$.

$$
\begin{aligned}
\mathcal{L} &= \frac{1}{(k+1)^n} \log\left(\prod_{i=-0}^{n} \left(\binom{n}{i} k^i\right)!\right) \\
&= \sum_{i=0}^{n} \frac{\log\left(\left(\binom{n}{i} k^i\right)!\right)}{(k+1)^n} \\
&= \sum_{i=0}^{n} \frac{\log\left(\sqrt{2\pi\binom{n}{i}k^i}\left(\frac{\binom{n}{i}k^i}{e}\right)^{\binom{n}{i}k^i}(1+o(1))\right)}{(k+1)^n} && (70)\\
&= \sum_{i=0}^{n} \frac{\binom{n}{i}k^i\left[\log\left(\binom{n}{i}k^i\right) - \log(e)\right]}{(k+1)^n} + o(1) && (71)\\
&= \sum_{i=0}^{n} \frac{\binom{n}{i}k^i \log\left(\binom{n}{i}k^i\right)}{(k+1)^n} - \log(e) + o(1). && (72)
\end{aligned}
$$

In (70) we use Stirling's Formula:

$$
n! \sim \sqrt{2\pi n}\left(\frac{n}{e}\right)^n (1+o(1)), \tag{73}
$$

to estimate the behavior of the logarithm of a factorial:

$$
\log(n!) = n \log n (1+o(1)).
$$

We must be careful in the implementation of Stirling's formula here. The $o(1)$ term in (73) goes to zero as n goes to infinity. To apply Stirling's to $\left(\binom{n}{i}k^i\right)!$ we must verify for all i that this factorial goes to infinity; each term has an acceptable rate of growth, except the first $i=0$, for which the term makes no contribution.

In (71), we employ the derivative of the binomial theorem:

$$
\sum_{i=0}^{n} \binom{n}{i} i x^{i-1} = n(1+x)^{n-1}.
$$

We continue the calculations from (72) by bounding the sum that appears there from both above and below. Beginning with the upper bound, we recall that for all $i = 1, \ldots, n$,

$$
\binom{n}{i} k^i \leq \binom{n}{\alpha n} k^{\alpha n} \text{ where } \alpha = \left\lfloor \frac{k}{k+1} \right\rfloor.
$$

We can pull the logarithm of this term out of the sum,

$$
\begin{aligned}
\mathcal{L} &= \sum_{i=0}^{n} \frac{\binom{n}{i} k^i \log\left(\binom{n}{i} k^i\right)}{(k+1)^n} - \log(e) + o(1) \\
&\leq \log\left(\binom{n}{\alpha n} k^{\alpha n}\right) - \log(e) + o(1) \\
&= \left[\frac{k \log k}{k+1} + h\left(\frac{k}{k+1}\right)\right] n - \log(e) + o(1) \\
&= n \log(k+1) - \log(e) + o(1).
\end{aligned}
$$

We proceed with the lower bound for (72).

$$
\begin{aligned}
\mathcal{L} &= \sum_{i=0}^{n} \frac{\binom{n}{i} k^i \log\left(\binom{n}{i} k^i\right)}{(k+1)^n} - \log(e) + o(1) \\
&= -\frac{1}{(k+1)^n} \sum_{i=0}^{n} \binom{n}{i} k^i \log\left(\frac{1}{\binom{n}{i} k^i}\right) - \log(e) + o(1) \\
&\geq -\log\left(\frac{1}{(k+1)^n}(n+1)\right) - \log(e) + o(1) \\
&\geq n \log(k+1) - \log(n+1) - \log(e) + o(1).
\end{aligned}
$$

Above we use Jensen's Inequality for the convex function $-\log(x)$. As stated in Proposition 4.2.1.

The asymptotics for $\mathcal{L}$ above show that the upper and lower bounds both have matching linear in n leading term. Proposition 4.2.1 will be verified once we show the additional term in the Brightwell-Tetali upper bound has smaller than linear order. The remaining term from (69) is:

$$
\begin{aligned}
\frac{1}{(k+1)^n} \log\left(\left(\frac{2e(K-1)N}{r}\right)^r\right) &= \frac{1}{(k+1)^n} \log\left(\left(\frac{2en(k+1)}{r}\right)^r\right) \\
&= \frac{r \log\left(\frac{(k+1)^n}{r}\right) + r \log(2en)}{(k+1)^n}.
\end{aligned}
\tag{74}
$$

To continue this expansion, we must calculate the quantity r for $F_{n,k}$, as described in (59):

$$
r = \sum_{i=1}^{n} \frac{r_{i-1}}{d_i} = \sum_{j=0}^{n-1} \frac{k^j \binom{n}{j}}{j+1} = \frac{(k+1)^{n+1} - 1}{k(n+1)} - \frac{k^n}{n+1} = \frac{(k+1)^{n+1} - k^{n+1} - 1}{k(n+1)}. \tag{75}
$$

The second sum is a standard equation whose closed form is given in [38] (page 34).

We will first examine the quantity

$$\begin{aligned}\frac{r}{(k+1)^n} &= \frac{(k+1)^{n+1} - k^{n+1} - 1}{k(n+1)(k+1)^n} \\ &= \frac{k+1}{k(n+1)} + \frac{\left(1 - \frac{1}{(k+1)}\right)^n}{n+1} - \frac{1}{k(n+1)(k+1)^n}.\end{aligned} \tag{76}$$

From this expansion for r we can see that the leading order term is $\frac{k+1}{k(n+1)}$. Since $1 \leq \frac{k+1}{k} \leq 2$ we can see that $\frac{r}{(k+1)^n} = \Theta\left(\frac{1}{n}\right)$. Returning to the expression at the end of (74), we see that we can write:

$$\begin{aligned}\frac{r \log\left(\frac{(k+1)^n}{r}\right) + r\log(2en)}{(k+1)^n} &= \frac{r}{(k+1)^n} \log\left(\frac{(k+1)^n}{r}\right) + \frac{r}{(k+1)^n}\log(2en) \\ &= \Theta\left(\frac{\log n}{n}\right).\end{aligned}$$

□

Instead of using Brightwell-Tetali bound in (69) we could have proceeded by using the Kleitman-Sha bound as stated in Theorem 4.1.3. Using this simpler bound we still get agreement in the first order term; unless we use better bounds on the sum appearing in (72) we will not see the additional improvement in the constant term between the Kleitman-Sha and Brightwell-Tetali bounds.

Recall that the mode of the rank sequence of $F_{n,k}$ occurs at rank $\alpha n = \lfloor \frac{k}{k+1} \rfloor n$. The logarithm of the *size* of this level has the same order as our leading term. Given a random linear extension of $F_{n,k}$, the leading term from Proposition 4.2.1 states that the majority of the information about this linear extension is conveyed by recording the positions of all elements from the largest level. It takes $\log\left(k^{\alpha n}\binom{n}{\alpha n}\right)$ bits to record these positions.

CHAPTER V

DEDEKIND TYPE PROBLEMS

Dedekind's problem, that of counting the number of antichains in the Boolean lattice, has historically generated great interest. The problem was first posed by Dedekind in 1897 [14]. There is no known closed form solution to this problem; According to the Online Encyclopedia of Integer Sequences [45], only the first 8 values are known exactly. Asymptotic bounds with varying degrees of accuracy have been given (see [29], [31], [39], and [26].)

In this chapter, we will present two extensions of the classical Dedekind's problem; we will be estimating the number of antichains in posets which are generalizations of the Boolean lattice: the poset of partially defined functions, $F_{n,k}$, and the chain product poset, $[t]^n$. These posets lack some of the 'nice' properties of the Boolean lattice, such as symmetry about a central rank and a single parameter which determines the poset. Additionally, $[t]^n$ is not ranked, so it lacks degree regularity within level sets, and $F_{n,k}$ lacks a lattice structure, as it has no unique maximal element. These difficulties prevent the use of the techniques of Saphozhenko [39] and Korshunov [31], which give asymptotics for the number of antichains itself. We will instead give estimates on the logarithm of the number of antichains, paralleling the information theoretical approaches utilized by Kahn [26] and Pippenger [36].

5.1 Antichains in the Function Poset

Recall that $F_{n,k}$ is the poset of partially defined functions from $[n]$ to $[k]$ ordered by extension. This is a generalization of the Boolean lattice, as $F_{n,1}$ is isomorphic to the Boolean Lattice $\mathfrak{B}_n$. $F_{n,2}$ is also closely related to $\mathfrak{B}_n$. If you reverse all of the edges in the Hasse diagram of $F_{n,2}$, (a process called *taking the dual of the poset*), you get the cubical lattice, Q_n. Q_n is a natural extension of the Boolean lattice, as it is formed by taking all of the lower-dimensional faces (i.e. individual points, edges, etc.) of $\mathfrak{B}_n$ and ordering them by inclusion. Since the number of antichains is preserved under the dual operation, substituting the value $k = 2$ into any bound for the number of antichains in $F_{n,k}$ also gives a bound for the number of antichains in Q_n.

Let $a(P)$ represent the number of antichains contained in a poset P. In this section, our goal is to give an estimate for $\log(a(F_{n,k}))$. The following theorem of Kahn [26], will play a key role:

Theorem 5.1.1 (Kahn) *Let P be a graded poset with levels $P_1, P_2, \ldots, P_m$, with $|P_m| \leq M$. Assume that there exists an $s \in \mathbb{N}$ (so that s is a uniform bound), with $s \geq d_{up}(v)$ and $s \leq d_{down}(v)$ for all $v \in P$, then*

$$a(P) \leq (m2^s - (m-1))^{\frac{M}{s}}. \tag{77}$$

We note that Kahn proved this theorem so he could bound the number of antichains in the Boolean Lattice $B_n \cong F_{n,1}$. However, the theorem cannot be applied directly, as neither $F_{n,k}$ nor B_n satisfy all of the hypothesis of the theorem. Kahn proves a technical lemma allowing him to apply his theorem indirectly to B_n; we will prove a similar technical lemma which allows us to extend his result to $F_{n,k}$.

The i^{th} level set in $F_{n,k}$ has size $|P_i| = \binom{n}{i}k^i$, for $i = 0, \ldots, n$. Let $s_{n,k}$ be the index of the largest level set in $F_{n,k}$. By maximizing this quantity in terms of i, we see that $s_{n,k} = \left\lfloor \frac{kn}{(k+1)} \right\rfloor$. Note that when $k = 1$, as in the case of $\mathfrak{B}_n$, we reclaim the well known fact that the largest level of the Boolean lattice has rank $\left\lfloor \frac{n}{2} \right\rfloor$. This definition leads directly to the lower bound in the following theorem, as any subset of the largest level set is itself an antichain.

Theorem 5.1.2 $2^{|P_{s_{n,k}}|} \leq a(F_{n,k}) \leq \left(s_{n,k}2^{s_{n,k}} - s_{n,k} - 1\right)^{\frac{|P_{s_{n,k}}|}{s_{n,k}}}$.

Calculating the quantity $\dfrac{\log(a(F_{n,k}))}{|P_{s_{n,k}}|}$ gives us a way of seeing how close the bounds are. As we can see, this result gives matching first order terms at the logarithmic level.

Corollary 5.1.3 *Recalling that* $s_{n,k} = \lfloor \frac{kn}{k+1} \rfloor$,

$$1 \leq \frac{\log(a(F_{n,k}))}{|P_{s_{n,k}}|} \leq 1 + O\left(\frac{\log(s_{n,k})}{s_{n,k}}\right).$$

$F_{n,k}$ does not satisfy the hypothesis for Kahn's theorem, as its rank-sequence is unimodal, and there is no uniform bound on degrees which applies to all levels. However, if we truncate $F_{n,k}$ at level $s_{n,k}$, the mode of the rank sequence, the truncated poset has a strictly increasing rank sequence and it satisfies that $d_{up}(x) \geq s_{n,k}$ for all $x \in P_1 \cup P_2 \ldots \cup P_{s_{n,k}-1}$ and $d_{down}(x) \leq s_{n,k}$ for all $x \in P_2 \cup P_3 \ldots \cup P_{s_{n,k}}$. Note that for $x \in P_i$, $d_{up}(x) = k(n-i)$ and $d_{down}(x) = i$; so $s_{n,k}$ appropriately plays the role of s in Theorem 5.1.1. Since we can use Kahn's theorem to count the number of antichains in the truncated poset, we seek a way to relate that number to the number of antichains in all of $F_{n,k}$.

Definition 5.1.4 *A poset Q is a* ***relaxation*** *of a poset P, if P and Q have the same groundset and $y \lessdot x \in Q \Rightarrow y \lessdot x \in P$.*

It is a simple consequence of the definition of relaxation that if Q is a relaxation of P, then $a(Q) \geq a(P)$. The Hasse diagram of P may contain more edges than that of Q, but adding edges only decreases the number of possible antichains. In order to establish the upper bound in Theorem 5.1.2, we seek a poset which satisfies the hypothesis of Kahn's theorem and contains a relaxation of $F_{n,k}$ as a subposet. Kahn has already constructed such a poset for $F_{n,1} \cong B_n$, and here we prove the following technical lemma for $k \geq 2$:

Lemma 5.1.5 *Fix n and $k \in \mathbb{N}$ with $k \geq 2$; there exists a graded poset $A_{n,k}$ ranked by $\{0, 1, \ldots, n\}$ which:*

- *Contains a relaxation of $F_{n,k}$;*
- *Satisfies $d_{up}(x) \geq s_{n,k}$ for all $x \in A_1 \cup A_2, \ldots \cup A_{n-1}$ and $d_{down}(x) \leq s_{n,k}$ for all $x \in A_2 \cup A_3, \ldots \cup A_n$;*
- *Satisfies $|A_i| \leq |P_{s_{n,k}}| = \binom{n}{s_{n,k}} (k^{s_{n,k}})$ for all level sets $\{A_i\}_{i=0}^{n}$.*

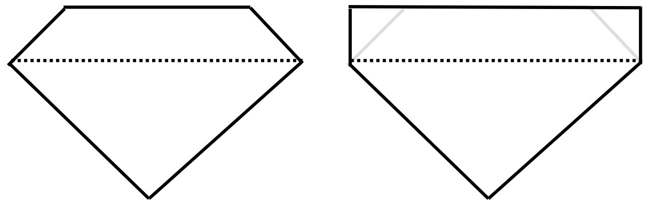

Figure 6: Heuristic drawings showing the shape of $F_{n,k}$ and $A_{n,k}$

Proof. Take the Hasse diagram for $F_{n,k}$, with $k \geq 2$, and consider it for the moment as a graph. Our objective is to manipulate it to form the Hasse diagram of an appropriate poset $A_{n,k}$, and then take the transitive closure of the covering relations represented as edges of this Hasse diagram to form the desired poset. We do not need to make any modifications below level $s_{n,k}$, as these vertices already satisfy the degree bounds, and the ranks are increasing up to level $s_{n,k}$. We follow three steps to transform the Hasse diagram, and then justify that these steps can indeed be carried out.

1. for $i > s_{n,k}$, remove $i - s_{n,k}$ down edges from each vertex on level i. These edges can be chosen arbitrarily. This gives us $d_{down}(y) = s_{n,k}$ for every vertex $y \in P_{s_{n,k}} \cup \ldots \cup P_n$.

2. for $i > s_{n,k}$, add vertices so there are $k^{s_{n,k}}\binom{n}{s_{n,k}}$ vertices on level i. Let us call the number of vertices needed on level i, j_i, and note that $j_i = k^{s_{n,k}}\binom{n}{s_{n,k}} - k^i\binom{k}{i}$. This insures that the rank sequence of $A_{n,k}$ will be increasing.

3. for $i \geq s_{n,k}$, add edges between level i and $i+1$ to ensure that both the up and down degrees of every vertex above level $s_{n,k}$ are exactly equal to $s_{n,k}$, while maintaining that $F_{n,k}$ is contained as a relaxation.

The feasibility of the first two steps is clear. However in the third step, we need to verify that we can add enough edges to satisfy the degree requirements without adding additional edges between vertices in the original $F_{n,k}$ ground set to guarantee that our new graph will contain a relaxation of $F_{n,k}$.

After the first step, counting edges by their top endpoint, we see that there are $s_{n,k}\left(k^{i+1}\binom{n}{i+1}\right)$ edges remaining between levels i and $i+1$. Since we know vertices in level i require up degree $s_{n,k}$, we can see that the number of edges which we need to add between level i and $i+1$ is $s_{n,k}\left(k^i\binom{n}{i}\right) - s_{n,k}\left(k^{i+1}\binom{n}{i+1}\right)$. We need to be able to add all of these edges between vertices in $F_{n,k}$ on level i and new vertices on level $i+1$ in order to increase the up degree for vertices in level i to at least $s_{n,k}$.

Since the down degree of new vertices on level $i+1$ needs to be $s_{n,k}$, we can use all of the available edges from $L = k^i\binom{n}{i} - k^{i+1}\binom{n}{i+1}$ new level $i+1$ vertices to supplement the up degree of original vertices from level i. This leaves $k^{s_{n,k}}\binom{n}{s_{n,k}} - k^i\binom{n}{i}$ vertices on level $i+1$ which currently have degree 0 (Notice that this is exactly the number of added vertices on level i). At this point, all vertices have the correct degree except these new vertices which remain isolated on level $i+1$ and all of the added vertices on level i which have not been modified in the above procedure (all of them still have up degree 0 at this point).

We can label these isolated vertices in levels i and $i+1$ with labels $1 \ldots j_i$ and $1 \ldots j_{i+1} - L$ respectively. Now we add edges so that every vertex x_l on the bottom gets connected with each vertex on the top y_k for $k \in \{l, l+1, \ldots, l+s_{n,k}\}$ where sums are evaluated modulo j_i. Since $s_{n,k} \le j_i$ this process creates a simple graph, called this graph the circulant bipartite graph with degree $s_{n,k}$.

This now determines the complete Hasse diagram for $A_{n,k}$ by giving all of its covering relations. Taking the transitive closure of these relations gives us a poset satisfying the hypothesis of Theorem 5.1.1. □

Proof of Theorem 5.1.2 follows by applying Theorem 5.1.1 on $A_{n,k}$ to give an upper bound for $a(A_{n,k})$, noting that $a(F_{n,k}) \le a(A_{n,k})$ since we ensured that it contained a relaxation of $F_{n,k}$. We use the values $m = s_{n,k}$ and $M = k^{s_{n,k}}\binom{n}{s_{n,k}}$.

Observation 5.1.6 *We use only a uniform bound on the degrees. We currently do not know if there is a proof, perhaps giving a tighter bound, which allows us to use the up and down degree information as it varies over the level sets.*

Besides being of intrinsic interest, antichains also have a connection to coding theory.

Definition 5.1.7 *An* ***error-detecting code*** *is a set of codewords, when sent by an encoder* *through a channel to receiver, the receiver can determine if a single error has occurred.*

The sender and receiver have a prearranged set of codewords. If the receiver received a string which is not one of the words on the list, the receiver is said to have 'detected an error' and can ask for retransmission. In a typical fixed length coding setting, we send length n codewords with entries chosen from an alphabet of size k. (For general coding theory definitions see [32]).

An antichain in the Boolean Lattice, B_n (equivalently $F_{n,1}$) is a single error-detecting code with the alphabet $\{0, 1\}$. This follows from the fact that an antichain must have Hamming distance at least 2. Not all single error-detecting codes are antichains, since a single error-deteching code allows codewords to be related as long as they have Hamming distance greater than 1. Nevertheless, generating a family of antichains also generates a family of error detecting codes.

In the standard coding theoretical channel, we assume that errors can occur symmetrically, any letter of the alphabet can be received incorrectly as any other letter. Another kind of channel is called an erasure channel. In this type of channel, there is no error when letters are sent, however there is a chance that a bit of information will be erased. We can think of the erasure character as another element in the alphabet, as long as we can induce error in the sending mechanism as well as random eraure error in the channel. We can model this situation with the following generalization: an endcoder sends length n codewords consisting of an alphabet of $k + 2$ letters over a channel which has the property that it receives characters $\{0, \ldots, k\}$ correctly as transmitted each time, but when character $k + 1$ is transmitted, it could actually represent any of the other letters. We could avoid using the last character, and have a code which is vacuously error detecting, by virtue of the fact that the channel does not make errors when using the symbols $\{0, \ldots, k\}$. However, we could expoit the unique property of our channel to find more codes, as an antichain in $F_{n,k}$ is an error-detecting code in this universe. For instance, there are significantly more antichains in $F_{n,2}$ than in $\mathfrak{B}_n$, giving a way to generate additional codes in this specific coding environment.

5.2 Antichains in the Chain Product Poset

Each point in $[t]^n$ is represented by an n-tuple, $(x_1, x_2, \dots, x_n)$, where $0 \leq x_i \leq t-1$ for all $i = 1, \dots, n$. For $y = (y_1, y_2, \dots, y_n)$ let $x \prec y$ if and only if $x_i \leq y_i$ for all $1 \leq i \leq n$. Notice that the Boolean lattice $\mathfrak{B}_n$ is isomorphic to $[2]^n$. This is a ranked poset, as we can assign $rank(x) = \sum_{i=1}^{n} x_i$. However the poset is not graded, as the degree of a vertex does not depend only on its rank. The up degree of a vertex is equal to the number of positions where it takes a value less than n, and the down degree is equal to its number of nonzero entries. Since the poset is not graded, we cannot use Kahn's theorem as presented in the previous section to enumerate the antichains in $[t]^n$. We will instead appeal to a method which looks at how chains and antichains interact.

Let N be the size of the middle layer of the chain product $[t]^n$. It was recently shown by Mattner and Roos [34] that

$$N = t^n \sqrt{\frac{6}{\pi(t^2-1)n}} \left(1 + o(1)\right). \tag{78}$$

This gives an easy lower bound for the number of antichains contained in $[t]^n$, as each subset of the middle layer is in itself an antichain. We will show, using an information theoretic technique, that 2^N asymptotically approximates the number of antichains in $[t]^n$. More precisely,

Theorem 5.2.1 *Let $a([t]^n)$ be the number of antichains in $[t]^n$, then for $n \geq 4$, $t = o(n^\epsilon)$ with $0 < \epsilon \leq \frac{1}{8}$, we can say that:*

$$N \leq \log(a([t]^n)) \leq N \left(1 + \frac{11t^2(\log n)^{\frac{3}{2}}}{n^{\frac{1}{4}}}\right).$$

The $t = o(n^\epsilon)$, $0 < \epsilon < \frac{1}{8}$ hypothesis is necessary to ensure that this theorem gives matching first order terms at the logarithmic level. In the proof itself we use a weaker hypothesis on the relationship of t and n– the theorem remains true when $t = \omega(n^{\frac{1}{8}})$, however in this case the result gives us no useful information.

We follow closely a method used by Pippenger [36], which he used to show a similar result for the number of antichains contained in the Boolean lattice:

Theorem 5.2.2 *Let $a(\mathfrak{B}_n)$ be the number of antichains in the Boolean lattice. Then:*

$$\binom{n}{\lfloor \frac{n}{2} \rfloor} \leq \log(a(\mathfrak{B}_n)) \leq \binom{n}{\lfloor \frac{n}{2} \rfloor}\left(1 + \frac{(\log n)^{\frac{3}{4}}}{n^{\frac{1}{4}}}\right). \tag{79}$$

We know that Pippenger's result does not give the best known asymptotics for the number of antichains in the Boolean Lattice. In particular, the second order term is far from what we know to be the truth in this case, which was discovered by Korshunov [31], who developed asymptotics for the function $a(\mathfrak{B}_n)$ directly. The detailed case analysis involved in his argument is very precise, and difficult to reproduce in a ranked poset which is not bi-regular. A result of this strength is seemingly out of reach with current entropy techniques. We intend Theorem 5.2.1 as a preliminary result, to establish that N gives the correct first order term at the logarithmic level for the number of antichains in $[t]^n$.

For our information theoretic approach, we will use two functions extensively, $H(X)$, the entropy of a random variable (see [2] for information on entropy), and $h_1(p)$, the truncated binary entropy function. We define

$$h_1(q) = \begin{cases} -q\log_2 q - (1-q)\log_2(1-q) & \text{if } 0 \leq q \leq \frac{1}{2} \\ 1 & \text{if } \frac{1}{2} \leq q \leq 1. \end{cases} \tag{80}$$

A function f defined on a poset $(P, \preceq)$ is monotone if $x \preceq y$ implies that $f(x) \leq f(y)$. We note that there is a one-to-one correspondence between antichains and monotone Boolean functions in a ranked poset. We can represent an antichain, $\mathcal{A}$, with a Boolean function g, by taking $g(a) = 1$ for all $a \in \mathcal{A}$, and $g(b) = 0$ otherwise. From each such function g, we can form the associated monotone Boolean function f, by "closing g upwards," i.e. by setting $f(z) = 1$ if $x \preceq z$ for some x for which $g(x) = 1$. Similarly, given a monotone Boolean function f, note that the set of minimal elements x, for which $f(x) = 1$, defines an antichain. This correspondence allows us to count antichains by counting monotone (Boolean) functions in $[t]^n$.

We will use the following lemma from [36] in the course of our argument.

Lemma 5.2.3 *Suppose the random variable K takes values in $\{0, 1, \ldots, n\}$, and for some $k \geq 1$ and $0 \leq q \leq 1$,*

$$\mathbb{P}(K \geq k) \leq q,$$

Then $H(K) \leq h_1(q) + \log k + q \log n$.

Note that $[t]^n$ is a Sperner poset, so the largest antichain in $[t]^n$ is a level set. This follows from the fact that it is the product of chains, which are Sperner. Together with Dilworth's Theorem, which states that there is a chain partition whose size is equal to the size of the largest antichain, we are guaranteed that there is a chain partition of $[t]^n$ whose size is exactly N. (See [16] for additional details on Sperner posets and chain partitions). We will fix one such chain partition for the remainder of our argument. We will enumerate the elements of this chain partition with $\{C_1, C_2, \ldots, C_N\}$. For a fixed monotone function, g, there is a unique point in each chain where the values of the function change from zero to one. For each chain C_j in the partition, we can define a parameter to capture this information,

$$\gamma_j(g) = \#\{x \in C_j : g(x) = 1\}\,.$$

Our strategy for counting the number of monotone functions will hinge on this property. Let f be a monotone Boolean function, chosen uniformly from the set of all monotone Boolean functions. We know from a basic property of entropy that

$$H(f) = \log(a([t]^n)).$$

We will define a pair of variables $(\hat{\delta}, \tilde{\delta})$, which will in turn determine f. This will allow us to use the subadditivity of entropy again to give an upper bound on $H(f)$:

$$H(f) \leq H(\hat{\delta}, \tilde{\delta}) \leq H(\hat{\delta}) + H(\tilde{\delta}). \tag{81}$$

To define $\tilde{\delta}$ we first need another description. For each point $x \in [t]^n$, and $\ell = 0, \ldots, t-1$, define $d_\ell(x)$ to be the number of coordinates of x which take value ℓ. Let $d_{down}(x)$ be the down degree of x, the number of neighbors of x on the immediately preceding level. We can think of $d_\ell(x)$ as the ℓ^{th} down degree of x, as $\sum_{\ell=1}^{t-1} d_\ell(x) = d_{down}(x)$. We will call a point in $[t]^n$ *low* if $d_j(x) < \frac{n}{2t}$ for some $1 \leq j \leq t-1$. The word low is a vestigial artifact of Pippenger's proof, where he called points low if $d_1(x) \leq \frac{n}{4}$. These points directly correspond to all points in the Boolean lattice which occur on the lowest $\frac{n}{4}$ levels. In the context that we use the word low, it is important to note that it is possible to have two points in the same level set where one is low and the other is not, so low is not purely rank dependent. *This is indeed a new ingredient of our proof.* We will call a chain low if it contains a low point.

We will take f and selectively "forget" information from some of the chains in the chain partition. To be more precise, let $v_1, v_2, \ldots, v_N$ be independent random variables assigned to each chain in the chain partition. Let $p = \frac{(\log n)^{\frac{1}{2}}}{n^{\frac{1}{4}}}$. With each chain C_j, we associate a variable as follows:

$$v_j = \begin{cases} 1 \text{ if } C_j \text{ is low} \\ 1 \text{ with probability } p \text{ for } C_j \text{ not low} \\ 0 \text{ with probability } 1-p \text{ for } C_j \text{ not low} \end{cases} \tag{82}$$

From f we form the function $\tilde{\delta} = (\tilde{\delta}_1, \tilde{\delta}_2, \ldots, \tilde{\delta}_N)$, by taking $\delta_j = \gamma_j(f)v_j$. Now $\tilde{\delta}$ gives us enough information to reconstruct f on all low chains and on any chain C_j, with $v_j = 1$. Let $\tilde{f}$ be the smallest monotone function which is consistent with $\tilde{\delta}$, i.e. the smallest function so that $\gamma_j(\tilde{f}) \geq \tilde{\delta}_j$, for all $1 \leq j \leq N$. We record the missing information about f in a variable $\hat{\delta} = (\hat{\delta}_1, \hat{\delta}_2, \ldots, \hat{\delta}_N)$ where $\hat{\delta}_j = \gamma_j(f) - \gamma_j(\tilde{f})$. From these definitions, it is easy to see that f is determined by $\delta = (\tilde{\delta}, \hat{\delta})$; indeed, $\tilde{\delta}$ contains the information about g on all chains which are not forgotten, and we can reclaim information about the rest of the chains using the information from $\hat{\delta}$.

We now want to bound $H(\tilde{\delta})$. First we use the subadditivity of entropy to say that:

$$H(\tilde{\delta}) \leq \sum_{j=1}^{N} H(\tilde{\delta}_j). \tag{83}$$

For each fixed j, we now bound $H(\tilde{\delta}_j)$ using Lemma 5.2.3. Observe that $\tilde{\delta}_j \geq 1$ only if $v_j = 1$. We can use $\mathbb{P}(v_j = 1)$ in place of q in the lemma, so that $\mathbb{P}(\tilde{\delta}_j \geq 1) \leq \mathbb{P}(v_j = 1)$ implies that $H(\tilde{\delta}_j) \leq h_1(q) + q \log n$. If C_j is low then $v_j = 1$ with probability 1. Then $\tilde{\delta}_j = \gamma_j$, a random variable taking either value 0 or $[1 \ldots n]$, since this records how many ones are in our low chain. Therefore if C_j is low, $H(\tilde{\delta}_j) \leq 1 + \log(n)$. If C_j is not low, we can use Lemma 5.2.3 with $q = p$. Letting M be the number of low chains, we can separate the terms of the sum as follows:

$$H(\tilde{\delta}) \leq M(1+\log n) + (N-M)(h_1(p) + p\log n). \tag{84}$$

To proceed, we need to give a bound on M, the number of low chains. We can bound this from above by the number of low points. In order for a point x to be low, we need there to be some j for which $d_j(x) \leq \frac{n}{2t}$. We can think of each coordinate x_i, $i = 1, \ldots, n$, as a uniform random variable chosen from $\{0, 1, \ldots, t-1\}$. Then $\mathbb{P}(x_i = j) = \frac{1}{t}$, so the expected value for $d_j(x) = \frac{n}{t}$. Using a version of Chernoff's inequality from [24], we see that $\mathbb{P}(d_j(x) \leq \frac{n}{2t}) \leq \exp(-\frac{(\frac{n}{2t})^2}{2\frac{n}{t}})$. Since this is for a single value of j, we multiply by both t and the total number of points, to see that there are at most $t^{n+1}\exp(-\frac{n}{8t})$ low points.

Plugging this estimate and the value for p back into (84) and recalling the value of N from (78), we see that:

$$\begin{aligned} H(\tilde{\delta}) &\leq t^{n+1}e^{-\frac{n}{8t}}(1+\log n) + N\left(\frac{3(\log n)^{\frac{3}{2}}}{n^{\frac{1}{4}}}\right) \\ &\leq N\left(\frac{4(\log n)^{\frac{3}{2}}}{n^{\frac{1}{4}}}\right). \end{aligned} \tag{85}$$

In the first inequality above, we use that for p small, $h_1(p) \leq 2p\log(1/p)$. In the second inequality we use the fact that $t = o(n^{\epsilon})$ (Though for this calculation it suffices for $t \leq \frac{n}{16\log n}$).).

Now we need to bound $H(\hat{\delta})$. Working with $\hat{\delta}$, we want to explore the possible discrepancy between f and $\tilde{f}$. We will call a chain C_j bad if $\hat{\delta}_j \geq 2$. Again, the first property we use is the subadditivity of entropy to say that:

$$H(\hat{\delta}) \leq \sum_{i=1}^{N} H(\hat{\delta}_i).$$

We can again apply Lemma 5.2.3, now allowing $k = 2$ and setting $q_j = \mathbb{P}(\hat{\delta}_j \geq 2)$. Let $Q = \sum_{j=1}^{N} q_j$, noting that this is the expected number of bad chains. Then we can continue to say that:

$$
\begin{aligned}
H(\hat{\delta}) &\leq \sum_{j=1}^{N}(h_1(q_j)+1+q_j\log n) \\
&\leq Nh_1\left(\frac{Q}{N}\right)+N+Q\log n \\
&= N\left(1+h_1\left(\frac{Q}{N}\right)+\frac{Q}{N}\log n\right).
\end{aligned} \tag{86}
$$

where we use the concavity of entropy in the second inequality.

We will break into two cases to find a bound on Q. A point x in $[t]^n$ is called bad if (i) x is not low, (ii) the chain C_j which contains x also contains some y so that y is a neighbor of x on the immediately preceding level with $f(y)=1$, and (iii) $\tilde{f}(x)=0$. Each one of x's immediately preceding neighbors arises by decreasing one of the nonzero coordinates of x by one. If y differs from x in a coordinate where the value of x is k, we will refer to y as a k-neighbor of x. We classify bad points into two groups, points which are bad because they have many such k-neighbors for some $k\geq 1$, and points which have relatively few k-neighbors for some $1\leq k\leq t-1$. To formalize this, let $s=n^{\frac{1}{4}}(\log n)^{\frac{1}{2}}$. We will call a point x heavy if for any k, x has more than s k-neighbors. In order for a heavy x to be bad, we need each of the (at least s) chains containing k-neighbors y with $f(y)=1$ to be assigned a γ value of 0. This happens with probability $(1-p)$ for each chain. Using $p=\dfrac{(\log n)^{\frac{1}{2}}}{n^{\frac{1}{4}}}$ allows the following calculation:

$$\mathbb{P}(x \text{ is heavy and bad}) \leq (t-1)(1-p)^s \leq (t-1)e^{-ps} \leq \frac{t}{n}.$$

The factor of $(t-1)$ exists because there are $t-1$ different k values for which x can be heavy. If x is not heavy, it means that for every k, the number of k-neighbors of x is less than s. We apply the group Sym(n), the group of all permutations on $[n]$. The subgroup Stab(x), which fixes x acts transitively on each collection of k-neighbors. Let y be a k-neighbor of x so that $f(y)=1$. We can average over the whole orbit of y, since x is only bad if the chain containing it also contains y. Then the probability that x is bad but not heavy is bounded by $\frac{s}{\frac{n}{2t}}=\frac{2ts}{n}=\frac{2t(\log n)^{\frac{1}{2}}}{n^{\frac{3}{4}}}$.

Combining these two estimates, we see that for $n\geq 4$,

$$\mathbb{P}(x \text{ is bad}) \leq \max\left(\frac{t}{n},\frac{2t(\log n)^{\frac{1}{2}}}{n^{\frac{3}{4}}}\right)=\frac{2t(\log n)^{\frac{1}{2}}}{n^{\frac{3}{4}}}.$$

Therefore $Q = \mathbb{E}(\text{bad points}) \leq t^n \frac{2t(\log n)^{\frac{1}{2}}}{n^{\frac{3}{4}}}$. We use bounds for Q and N in (86) to see that:

$$\begin{aligned} H(\hat{\delta}) &\leq N\left(1 + h_1\left(\frac{Q}{N}\right) + \frac{Q}{N}\log n\right) \\ &\leq N + \left(\sqrt{\frac{\pi}{24}}\right)\frac{t^2(\log n)^{\frac{3}{2}}}{n^{\frac{1}{4}}} + \left(2\sqrt{\frac{\pi}{6}}\right)\frac{t^2(\log n)^{\frac{3}{2}}}{n^{\frac{1}{4}}} \\ &\leq N + \frac{7t^2(\log n)^{\frac{3}{2}}\log t}{n^{\frac{1}{4}}}. \end{aligned} \tag{87}$$

Now we are ready to give estimates for both parts of (81), to finally conclude that:

$$\begin{aligned} H(f) &\leq H(\tilde{\delta}) + H(\hat{\delta}) \\ &\leq N + \frac{7t^2(\log n)^{\frac{3}{2}}\log t}{n^{\frac{1}{4}}} + N\left(\frac{4(\log n)^{\frac{3}{2}}}{n^{\frac{1}{4}}}\right) \\ &\leq N\left(1 + \frac{11t^2(\log n)^{\frac{3}{2}}\log t}{n^{\frac{1}{4}}}\right) \end{aligned} \tag{88}$$

as desired.

5.3 Butterfly-free Functions

The previous two sections have expanded the original scope of Dedekind's Problem by counting antichains in posets which are generalizations of the Boolean Lattice. In this area, another direction to explore is counting the number of structures which are in some way a generalization of antichains in the Boolean Lattice itself. We will introduce a more complex structure whose definition, similar to that of antichains, relies upon inclusion. However, unlike antichains, the definition relies on the relationship between multiple sets.

Definition 5.3.1 *A collection of four distinct sets, $A, B, C, D \in \mathfrak{B}_n$ is called a* ***butterfly*** *if $A \cup B \subseteq C \cap D$.*

The name butterfly was coined for this structure by Jerry Griggs, because a non-degenerate arrangement of these sets evokes the image of a butterfly. However, as you can see from Figure 5.3, not all butterflies possess this evocative shape. The definition of butterfly-free rules out not necessarily induced structures; for instance, chains of length four are not permitted, as it would provide a butterfly with the union of the lower two

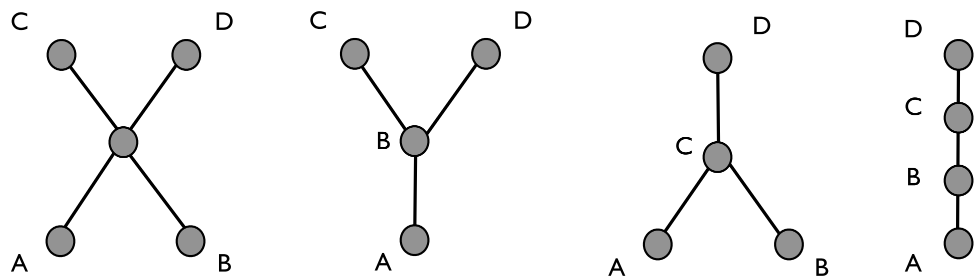

Figure 7: Some possible arrangements of sets forming Butterflies

sets being contained in the intersection of the higher two. Additionally, the difference in rank between the sets in a butterfly is not constrained, so the shape does not have to have vertical symmetry.

We can see that not only are butterfly-free families a generalization of inclusion-free families (i.e. antichains), they are also a generalization of union-free families. We will call a set system in B_n "butterfly-free" (BFF) if it contains no butterflies. We will use the standard correspondence between set systems and Boolean functions to also refer to *butterfly-free functions.* Let $\mathfrak{F}$ be the family of all BFF Boolean set systems. Following the intent of Dedekind's Problem, we ask 'how many butterfly-free families are there?' or 'what are bounds on $|\mathfrak{F}|$?'

Since the property of being butterfly-free is closed under taking subfamilies, the collection of all subsets of the largest BFF family gives us a trivial lower bound on $|\mathfrak{F}|$. Let F be a butterfly-free family. De Bonis *et al.* [13] show that

$$|F| \leq \binom{n}{\lfloor \frac{n}{2} \rfloor} + \binom{n}{\lceil \frac{n}{2} \rceil}.$$

It is clear that any family which is restricted to two adjacent layers of the Boolean Lattice is BFF, as the sets of a butterfly must span a minimum of three levels. The middle two layers of B_n form a BFF family of maximum size, so their bound for the size of the largest BFF family is tight. This observation gives the lower bound in the theorem below:

Theorem 5.3.2

$$\binom{n}{\lfloor \frac{n}{2} \rfloor} + \binom{n}{\lceil \frac{n}{2} \rceil} \leq \log(|\mathfrak{F}|) \leq 4\sqrt{2}\ (\binom{n}{\lfloor \frac{n}{2} \rfloor})(1 + o(1)).$$

Proof. The proof is an adaption of a proof by Burosch et al. [10]. (Their proof was originally used to count the number of database closures, which are in one-to-one correspondence to union-free families.) Let $\mathbb{F}$ be a butterfly-free family on B_n. Partition the ground set $[n]$ arbitrarily into two classes X_1 and X_2, of as equal size as possible. We will define four families of sets:

$$F_1 = \{F \in \mathbb{F} | \nexists G_1 \in \mathbb{F} \text{ so that } F \cap X_1 = G_1 \cap X_1 \text{ and } F \cap X_2 \subsetneq G_1 \cap X_2\}$$
$$F_2 = \{F \in \mathbb{F} | \nexists G_2 \in \mathbb{F} \text{ so that } F \cap X_2 = G_2 \cap X_2 \text{ and } F \cap X_1 \subsetneq G_2 \cap X_1\}$$
$$F_3 = \{F \in \mathbb{F} | \nexists G_3 \in \mathbb{F} \text{ so that } G_3 \cap X_1 = F \cap X_1 \text{ and } G_3 \cap X_2 \subsetneq F \cap X_2\}$$
$$F_4 = \{F \in \mathbb{F} | \nexists G_4 \in \mathbb{F} \text{ so that } G_4 \cap X_2 = F \cap X_2 \text{ and } G_4 \cap X_1 \subsetneq F \cap X_1\}.$$

We will proceed with two lemmas:

Lemma 5.3.3 $\mathbb{F} = F_1 \cup F_2 \cup F_3 \cup F_4$.

Suppose not, then there exists a set $F \in \mathbb{F}$ which is not in any F_i; this means that for each $i = 1, \ldots, 4$ there exists $G_i \in \mathbb{F}$ satisfying the above criteria. However, this creates a contradiction, since $\mathbb{F}$ was taken to be a BFF family. We can easily see that the G_i's are all distinct members of $\mathbb{F}$ and that $G_1 \cap G_2 = F$ and $G_3 \cup G_4 = F$, so that we have $G_3 \cup G_4 = G_1 \cap G_2$, creating a butterfly in $\mathbb{F}$.

□

Fix $j \in \{1, 2\}$, and a set $A \in X_j$. We define a collection of sets:

$$F_i(A) = \{B \in X_{j+1\,(mod\,2)} \,|\, A \cup B \in \mathbb{F}\}.$$

Lemma 5.3.4 *Fix $j \in \{1, 2\}$, for any set $A \in X_j$ and for $i = 1 \ldots 4$, collections of sets $F_i(A)$ form an antichain in the m dimensional Boolean Lattice for $m = |X_{j+1\,(mod\,2)}|$.*

Because of the symmetry in the definitions of the sets, it is sufficient to show for any A that $F_1(A)$ is inclusion-free. Suppose there exist sets $B_1 \subset B_2 \in F_1(A)$. For ease of notation, let us define $B'_k = A \cup B_k$ for $k = 1, 2$. Then from the definition of $F_1(A)$ we see that $B_1' \cap X_1 = B_2' \cap X_1 = A$. Since B_1 and B_2 are distinct, $B_1 = B_1' \cap X_2 \subsetneq B_2' \cap X_2$, therefore B_1 cannot be a member of $F_1(A)$ since the definition of $F_1(A)$ specifically state that no such set exists. This shows that $F_1(A)$ is an antichain, its ground set is X_2, so it is an antichain which resides in m dimensional Boolean Lattice for $m = |X_{j+1\,(mod\,2)}|$. $\square$

To proceed with the proof, we need merely to bound the possible number of values available for each F_i to take as $\mathbb{F}$ ranges over $\mathfrak{F}$. Call this quantity $\mathfrak{F}_\mathfrak{i}$. We will use the same bound for all of the F_i, so we will give the calculation only for F_1. By the second lemma, we see that for each $A \in X_1$, $\mathfrak{F}_1(A)$ is a Boolean antichain on a ground set of size at most $\lceil \frac{n}{2} \rceil$. Using Kahn's bound from [26], we have that

$$|\mathfrak{F}_i(A)| \leq 2^{\binom{\lceil \frac{n}{2} \rceil}{\lfloor \frac{n}{4} \rfloor}(1+o(1))}.$$

In the worst case scenario, for each $A \in X_1$, the families $F_1(A)$ are chosen independently of each other. This gives the bound that:

$$|\mathfrak{F}_1| \leq \Pi_{A \subseteq X_1} |\mathfrak{F}_1(A)| \leq 2^{2^{\frac{n}{2}} \binom{\lceil \frac{n}{2} \rceil}{\lfloor \frac{n}{4} \rfloor}(1+o(1))} \leq 2^{\sqrt{2}\binom{n}{\lfloor \frac{n}{2} \rfloor}(1+o(1))}.$$

The fact that $\mathfrak{F} \subseteq \Pi \mathfrak{F}_\mathfrak{i}$ for $i = 1, \ldots, 4$ gives the desired result. $\square$

CHAPTER VI

FAMILIES WHICH EXCLUDE CHERRIES

6.1 Introduction and Motivation

We can think of building an antichain in a poset by adding elements successively; an element can be added only if no elements which it is related to are already included in the antichain. We can also build a family of elements by adding a new element only when no two of its downward neighbors are already included in the family. Families of this type have relevance in several contexts, such as union-free families and Boolean Horn functions. In this chapter, we give bounds for the maximum size of such a family in the case that our poset is regular and bipartite. In the final section, we introduce Horn functions on the Boolean lattice, and give an application of our result.

Having tight bounds on the asymptotic size of the largest union-free subset of $\mathfrak{B}_n$ (see Chapter 2) gives us a good lower bound on the number of union-free families of $\mathfrak{B}_n$. With hopes of improving the upper bound we investigate the number of cherry functions (defined below) in bipartite posets. Looking at a property restricted to bipartite posets and then extending the result to ranked partially ordered sets has been a successful tactic for counting the number of several other graph structures: graph homorphisms [20], linear extensions [8], and antichains [26]. Though in this chapter we provide results for bipartite posets, we have not been able to generalize to all ranked posets; we do interpret our results for relevance in the case of the Boolean Lattice. Compared with the task of counting antichains, it is harder to make the leap to ranked posets in the union-free case because of the property's reliance upon data from multiple neighbors; as it turns out, this is not an easy hurdle to circumvent. It is not clear the best way to extend the definition of cherry to multiple levels in a computationally feasible way. In the vein of Kahn's antichain proof [26], it seems likely there is an elegant induction and entropy based proof to extend the result from two level bipartite posets to ranked posets. Finding this proof is a direction of continued interest.

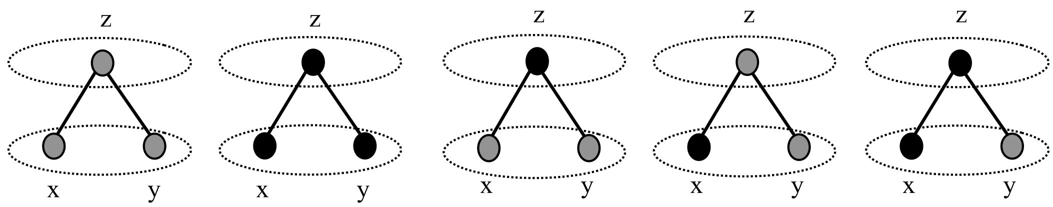

Figure 8: Acceptable configurations in a Cherry function. (Black vertices represent where the function takes the value 1.)

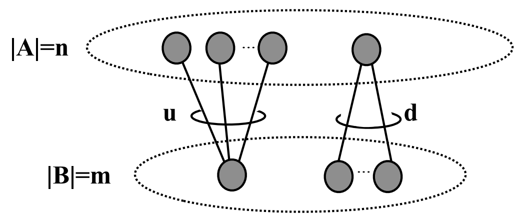

Figure 9: A (u,d)-poset

Definition 6.1.1 *On a two-level poset, a Boolean function g is said to have the **cherry property** if for all triples x, y, z so that $x \lessdot z$ and $y \lessdot z$ we have that $g(x) \wedge g(y) \leq g(z)$. For ease, we call such triples (x, y, z) **cherries**, and such functions **cherry functions**. Note this enforces exactly the property that if (x, y, z) is a cherry, for g to be a cherry function, if $g(x) = g(y) = 1$ then $g(z)$ must equal 1.*

Definition 6.1.2 *A **(u,d)-poset** is a bipartite regular poset with parts A and B, where elements of B have rank 0 and elements of A have rank 1. Additionally deg$(v) = u$ for all $v \in B$ and deg$(w) = d$ for all $w \in A$.*

Given a poset P, let $\mathfrak{C}(P)$ represent the number of cherry functions on P. In this section, we will show the following result:

Theorem 6.1.3 *Let P be a (u, d)-poset with n points on top and m points on the bottom. Then*

$$n + \frac{m}{ud - u + 1} \leq \log\left(\mathfrak{C}(P)\right) \leq n + \frac{m}{d}\log(d + 1).$$

Moreover, this upper bound is tight up to constants on the second order term, as it is achieved by $\frac{n}{d}$ disjoint copies of $K_{d,d}$.

Note that this result is the best possible up to logarithmic factors without additional restrictions on u and d, as it gives almost matching bounds if $u = 1$. The bounds are the least tight when $u = d$; For the regular case we will give a stronger lower bound for posets with an added co-degree hypothesis.

6.2 Upper Bounds for Bi-regular Two Level Posets

Let $P = A \cup B$, where $|A| = n$, $|B| = m$, with $d_{down}(x) = d$ for all $x \in A$ and $d_{up}(v) = u$ for all $v \in B$. Let $C(P)$ be the set of all cherry functions on P. So $|C(P)| = \mathfrak{C}(P)$.

Proposition 6.2.1 *Let f be a uniformly chosen cherry function on a (u,d)-poset, P. Then:*

$$H(f) \leq n + \frac{m}{d}\log(d+1).$$

Recalling from Proposition 1.4.5, $H(f) = \log(|C(P)|)$, and bounding $H(f)$ will allow us to give an upper bound for $\mathfrak{C}(P)$.

Since dn and um both count the number of edges in P, we must have $dn = um$. We may assume that $n > m$ (which implies that $u > v$) since there are more cherry functions on a bipartite graph when the top set is larger due to the asymmetry inherent in the definition of cherries. Proceeding, we use the following simplified notation: for $X \subseteq A \cup B$, we use f_X for $f|_X$ (i.e. f restricted to X), and similarly for $v \in A \cup B$ we use f_v for $f(v)$. If E is an event, $\overline{E}$ is its complement and $\mathbb{I}_E$ represents the indicator variable that the event E occurs.

Using the chain rule of entropy (Proposition 1.4.7) we can rewrite $H(f)$ as follows:

$$\begin{aligned} H(f) &= H(f_A|f_B) + H(f_B) \\ &\leq \sum_{v\in A}[H(f_v|f_{N(v)}) + H(f_B)]. \end{aligned} \tag{89}$$

To acquire the inequality $H(f_A|f_B) \leq \sum_{v\in A} H(f_v|f_{N(v)})$, we use subadditivity and the fact that f_v depends only on $f_{N(v)}$, not on all of B. Next we apply Shearer's lemma to the second term in (89) using the fact that each element of B is connected to u members of A:

$$
\begin{aligned}
H(f) &\leq \sum_{v\in A}\left[H(f_v|f_{N(v)}) + \frac{1}{u}H(f_{N(v)})\right] \\
&= \sum_{v\in A}\left[H(f_v|f_{N(v)}) + \frac{m}{nd}H(f_{N(v)})\right].
\end{aligned} \tag{90}
$$

To simplify the first term of this sum, we condition on an event which is related to the cherry property. For v ranging over A, let Q_v be the event that $\{\sum_{w\in N(v)} f_w \leq 1\}$. Note that if Q_v occurs, the value of f on v is not restricted by the values of f on $N(v)$. For each $v \in A$, let $q_v = \mathbb{P}(Q_v \text{ occurs})$. Conditioning the first term on Q_v gives us the expression:

$$H(f_v|f_{N(v)}) = q_v(H(f_v|\mathbb{I}_{Q_v} = 1)) + (1 - q_v)(H(f_v|\mathbb{I}_{Q_v} = 0)).$$

$H(f_v|\mathbb{I}_{Q_v} = 0)$ is zero, since f_v is determined if $\mathbb{I}_{Q_v} = 0$. The term $q_v(H(f_v|\mathbb{I}_{Q_v} = 1))$ can be greatly simplified with the observation that since $\mathbb{I}_{Q_v} = 1$, f takes the value 1 in $N(v)$ at most once. This means f_v can take either value 0 or 1. In this case, the function is unrestricted at v, so there must be a bijection between cherry functions taking the value 0 at v and those taking the value 1 at v. Since f_v takes two equally likely values, $H(f_v|\mathbb{I}_{Q_v} = 1) = 1$. Substituting this simplification back into (90), we are left with:

$$H(f) \leq \sum_{v\in A}\left[q_v + \frac{m}{nd}H(f_{N(v)})\right].$$

Now, to work with this second term which resulted from our use of Shearer's inequality, we again condition on $\mathbb{I}_{Q_v}$. Using the chain rule, we can write the joint entropy, $H(f_N(v), \mathbb{I}_{Q_v})$, by conditioning in two different ways.

$$
\begin{aligned}
H(f_{N(v)}, \mathbb{I}_{Q_v}) &= H(f_{N(v)}) + H(\mathbb{I}_{Q_v}|f_{N(v)}) \\
&= H(\mathbb{I}_{Q_v}) + H(f_{N(v)}|\mathbb{I}_{Q_v}).
\end{aligned}
$$

Since $\mathbb{I}_{Q_v}$ is completely determined by $f_{N(v)}$, we have that $H(\mathbb{I}_{Q_v}|f_{N(v)}) = 0$. Solving for $f_{N(v)}$ and substituting gives that:

$$
\begin{aligned}
H(f_{N(v)}) &= H(\mathbb{I}_{Q_v}) + H(f_{N(v)}|\mathbb{I}_{Q_v}) \\
&= H(q_v) + H(f_{N(v)}|\mathbb{I}_{Q_v}).
\end{aligned} \tag{91}
$$

The second equality above follows since $\mathbb{I}_{Q_v}$ is an indicator variable, $H(\mathbb{I}_{Q_v}) = H(q_v)$. (Recall that $H(x)$, for $0 \leq x \leq 1$, is the binary entropy function).

Using the estimate that for any random variable X, $H(X) \leq \log(|\text{range}(X)|)$, we observe that the range of $(f_{N(v)}|Q_v)$ is $d+1$, since we choose which, if any, of the d positions in $N(v)$ has the value one. Similarly the range of $(f_{N(v)}|\overline{Q_v})$ is $2^d - (d+1)$. Putting all of these pieces together, we see that:

$$\begin{aligned} H(f) &\leq \sum_{v \in A} q_v + \frac{m}{nd}\Big[H(q_v) + q_v \log(d+1) + (1-q_v)\log\Big(2^d - (d+1)\Big)\Big] \\ &\leq \sum_{v \in A} \frac{m}{nd}\Big[H(q_v) + q_v \log(d+1)\Big] + \Big(\frac{n-m}{n}\Big)q_v + \frac{m}{n}. \end{aligned} \tag{92}$$

Equality is acheived in the first equation above when we take P to be a single copy of $K_{d,u}$, showing that all of our equations are tight up to there. However, to make the bound more tractable, we make additional estimates.

We factor out a q_v term, and note that we can maximize the sum by assuming that each q_v is equal and then finding a universal bound for q_v. We set $q_v = q$ for all v to simplify notation. Continuing, we isolate terms involving q and take the derivative. Let the function $s(q)$ be all terms from (92) which depend on q. Then;

$$s(q) = \frac{m}{nd}H(q) + q\left(\frac{m\log(d+1) + nd - md}{nd}\right). \tag{93}$$

$$s'(q) = \frac{m}{nd}\left[\log\left(\frac{1-q}{q}\right) + \frac{m\log(d+1)}{nd} + \frac{n-m}{n}\right].$$

Setting this equal to zero and solving for q gives the optimal q value:

$$q = \frac{(d+1)2^{u-d}}{1+(d+1)2^{u-d}}.$$

We note that this value is very close to one, and that q is always less than 1, (since it is the probability that Q_v occurs). To simplify the final sum from (92), we will use $q_v = 1$, yielding the bound given in Proposition 6.2.1:

$$H(f) \leq n + \frac{m}{d}\log(d+1). \tag{94}$$

To get a result about regular two level posets, we let $n = m$ (implying that $d = u$). We get the estimate that for a d-regular two level poset, P on $2n$ vertices,

$$\log(\mathfrak{C}(P)) \leq n + \frac{n}{d}\log(d+1). \tag{95}$$

We will show this upper bound is tight for $\frac{n}{d}$ copies of $K_{d,d}$ by constructing a large set of cherry functions on this graph. For ease of notation, we refer to this graph as $Dk_{n,d}$. We can partition all cherry functions on $Dk_{n,d}$ into classes depending on how many points on top are "forced" when we only consider the function restricted to the bottom. To be more precise:

Definition 6.2.2 *A vertex v is* ***forced*** *by a function f if it has two or more down neighbors where f takes the value one.*

Note that this definition implies the function is forced to take the value zero at v. For a poset P, let $P(i)$ be the set of Cherry functions on P which have *exactly* i forced points. The size of any partition class is a lower bound for the total number of cherry functions. We will focus on the partition class $Dk_{n,d}\left(\frac{n}{2}\right)$. Notice that if a single copy of $K_{d,d}$ has no fixed points then there is at most one bottom point in that block which takes the value one. If any point in a copy of $K_{d,d}$ is fixed by f, then all of the top points will be fixed. Since the number of fixed points is always a multiple of d, the only way to fix $\frac{n}{2}$ points in $DK_{N,d}$ is to pick which of the $\frac{n}{2d}$ $K_{d,d}$ blocks will have fixed tops. This gives:

$$\begin{aligned}\left|Dk_{n,d}\left(\frac{n}{2}\right)\right| &\geq \binom{\frac{n}{d}}{\frac{n}{2d}}\left(2^d - d - 1\right)^{\frac{n}{2d}} 2^{\frac{n}{2}}\left((d+1)^{\frac{n}{2d}} - 1\right) \\ &\geq 2^n\, 2^{\frac{n}{d} + \log\left(\frac{n}{d}\right)}\, 2^{\frac{n}{2d}\log(d+1)} \\ &\geq 2^{n\left(1 + O\left(\frac{\log(d+1)}{d}\right)\right)}.\end{aligned} \tag{96}$$

The first two terms in the initial line count the choices for which blocks will have the tops fixed and the number of choices for the bottoms of those blocks. The remaining two terms are the number of choices for the values for the top and bottom points, respectively, in blocks which are not fixed.

6.3 Lower Bounds for (u, d)-posets

An easy lower bound for the number of cherry functions in a (u, d) poset notes that every antichain determines a unique cherry function. There are more than 2^{n+1} antichains, as this is how many we can acquire by choosing any subset of the bottom or any subset of the top. This gives a lower bound for both the number of antichains and the number of cherry functions. We can certainly do better than this.

A slightly improved lower bound can be found by adapting a technique from Sapozhenko, [39]: we take a fixed size subset of the bottom and calculate its possible influence on what can be added from the top. However, if the subset on the bottom has size larger than one the calculations become intractable because of uncertainty about how neighborhoods of the chosen vertices interact. The number of cherry functions on a d regular bipartite two level poset is clearly lower bounded if we count only cherry functions with a given number of ones on the bottom set.

Let $S \subset B$ with $|S| = \frac{n}{2^d}$. We define f by taking f to be one on S and zero on the rest of B. We look at the maximum number of elements that S can force to be one on the top. We get fthe largest set fixed if we assume all elements of S have disjoint neighborhoods. Additionally we enforce only the weaker requirement that an element in the top set is fixed if it has a single neighbor in S. This gives a vast undercounting of cherry functions with $\frac{n}{2^d}$ ones in the bottom, not just because of the weaker neighborhood requirement, but because if the neighborhoods are disjoint, no element in A will have two neighbors in S, and therefore will not be fixed in f.

Let the number of cherry functions with $\frac{n}{2^d}$ ones on the bottom be denoted by $\mathfrak{C}_{\frac{n}{2^d}}$, Then

$$\begin{aligned}
\log(\mathfrak{C}_{\frac{n}{2^d}}) &\geq \log\left(2^{n-d\frac{n}{2^d}}\right) + \log\left(\binom{n}{\frac{n}{2^d}}\right) \\
&= nH\left(\frac{1}{2^d}\right) + n - \frac{dn}{2^d} - \frac{\log n}{2} \\
&\geq n\left(1 + \frac{2^d-1}{2^{2d}} + \frac{2^d-1}{2^{3d+1}} - \frac{\log n}{2n}\right).
\end{aligned} \tag{97}$$

This gives us matching first order terms at the logarithmic level:

$$n + \frac{n}{2^{2d}}\left(2^d - \frac{1}{2^d} - o(n)\right) \leq \log(\mathfrak{C}(P)) \leq n + \frac{n}{d}\log(d).$$

As Graham Brightwell [7] pointed out, eliminating cherries favors selecting points from the top set over taking points from the bottom set. It is advantageous to approach this problem by taking an arbitrary subset of the top and then counting the number of compatible subsets of the bottom. If we fix an arbitrary subset of the top, we can use a simple greedy approach to choose a large set from the bottom which does not form any cherries when unioned with the top set. We will use this idea to show the following:

Proof of the lower bound from Theorem 6.1.3

Recall for P a (u, d)-poset, the theorem states that:

$$n + \frac{m}{ud - u + 1} \leq \log(\mathfrak{C}(P)). \tag{98}$$

Let S be an arbitrary subset of A. There are 2^n possible values for S. We want to choose a subset in B which forms no cherries with any possible S. Let M_1 be the set of points from B which unioned with any S forms a cherry free family. We initialize M_1 to be empty. Let M_2 be the set of points which are under consideration for membership in M_1. Initially we will take M_2 to be all of B. To proceed, arbitrarily choose some $v \in M_2$ and add it to M_1.

This v is connected to u elements on the top, and each of those elements are connected to $d-1$ elements on the bottom distinct from v. The second neighborhood of v, elements which have distance two from v, has size at most $u(d-1)$. (We don't get equality here since some of the elements in the bottom neighborhoods may overlap.) As long as we choose no elements from the second neighborhood of v, v will not be part of a cherry no matter which S is taken on top.

To continue, we remove v and its second neighborhood from M_2. We iterate, each time choosing a new v from M_2 to place into M_1 and then removing both v and its second neighborhood from M_2. We can see that each greedy choice removes at most $u(d-1)+1$ vertices from the m points initially in M_2; this process terminates after at least $\dfrac{m}{u(d-1)+1}$

steps. At each step we added one vertex to M_1, so $|M_1| \geq \dfrac{m}{u(d-1)+1}$. Any subset of M_1 yields a cherry function when unioned with any S, which yields the desired bound. $\square$

6.4 Co-Degree Hypotheses

Definition 6.4.1 *A poset P has **co-degree s** if given any two distinct points x, $y \in P$, $|N(x) \cap N(y)| \leq s$.*

Note that it is equivalent to say that a poset has co-degree 1 and that it is C_4-free.

Lemma 6.4.2 *Let $N_{n,d}$ be a two level poset with co-degree 1, where each point on the top has $d_{down} = d$ and the top set has cardinalilty n, then*

$$\log(\mathfrak{C}(N_{n,d})) \geq n\left(1 + \frac{d}{n} + O\left(\frac{\log n}{n}\right)\right).$$

Proof. Let the partition class of cherry functions with one point fixed be called $N_{n,d}(1)$. It is clear that $\mathfrak{C}(N_{n,d}) \geq |N_{n,d}(1)|$. We use the co-degree hypothesis to lower bound the size of this partition class. To calculate $|N_{n,d}(1)|$, we pick which vertex on top will be forced, we will call this forced point v_1. There are n choices for v_1. Since exactly one vertex is forced, we will have free choice for the function values on the remaining $(n-1)$ points in the top set. Looking at the neighborhood of v_1, we need to pick a subset of size at least 2 which takes the value 1, so that v_1 is truly forced. The co-degree condition guarantees that for $x, y \in N(v_1), N(x) \cap N(y) = \{v_1\}$. Therefore, setting any subset of v_1's neighbors to 1 will not force any other vertices on the top. This observation will give us a lower bound on $|N_{n,d}(1)|$. (We cannot assume this bound is tight, as there may be freedom to choose a subset $B \backslash N(v_1)$ which does not force any additional points.) We see that:

$$\begin{aligned}
\log(|N_{n,d}(1)|) &\geq \log\left(n2^{n-1}\left(\sum_{k=2}^{d}\binom{d}{k}\right)\right) \\
&\geq \log(n2^{n-1}(2^d - d - 1)) \\
&\geq n\left(1 + \frac{d}{n} + O\left(\frac{\log n}{n}\right)\right).
\end{aligned}$$

$\square$

For $N_{n,d}$ as described above which also is regular (the definition of $N_{n,d}$ does not require degree regularity of the bottom set which is important in the proof of the upper bound), we have that:

$$n\left(1+\frac{d}{n}+O\left(\frac{\log n}{n}\right)\right) \leq \log(\mathfrak{C}N_{n,d}) \leq n\Big(1+\frac{\log d}{d}\Big).$$

These bounds may seem hard to compare, but it is important that the constant subsumed in the Big O notation is actually quite small. Also note that $\ell(x) = \frac{\log x}{x}$ is decreasing in x. The co-degree condition guarantees that n is significantly larger than d, so $\frac{d}{n} + \frac{\log n}{n} \leq \frac{\log d}{d}$ will always hold.

Using similar ideas to those in the proof above, we can give a lower bound for the number of cherry function on (u, d) posets without requiring consistent co-degree restrictions. The construction involves finding the local optimal behavior for the neighborhood of each vertex. For each vertex $x \in A$, we define a parameter

$$s_x := \max_{U \subseteq N(x)} \{|U| : N(y) \cap N(z) = \{x\} \text{ for all } y, z \in U\}.$$

Let S_x be any subset of $N(x)$ with $|S_x| = s_x$. Define

$$f(S_x) = \max\left(1, \frac{m - s_x(u)(d-1)}{u(d-1)+1}\right).$$

Recall the definition of $P(1)$: the set of all cherry functions which have exactly one point fixed. Since this is a lower bound for $\mathfrak{C}(P)$ we make the following proposition:

Proposition 6.4.3 $\mathfrak{C}(P) \geq |P(1)| \geq \sum_{x\in A} 2^{n-1} \sum_{k=2}^{s_x} \binom{s_x}{k} f(S_x)$.

Proof. If we take a subset S_x which acheives s_x, then we can form a family of Cherry-free functions by setting $x \equiv 0$, any subset of $S_x \equiv 1$, and $N(x)\backslash S_x \equiv 0$. If we remove points in $N(N(S_x))$ and take a greedy selection of else remaining in B, we are guaranteed a free choice of $A\backslash\{x\}$. Each cherry function of this type has only x as a forced point, so is a member of $P(1)$. The quantity $\dfrac{m - s_x(u)(d-1)}{u(d-1)+1}$ in the definition of $f(S_x)$ is the number of points in $B\backslash N(N(S_x))$ which can be taken greedily from (as in the proof of (98)) while maintaining that $A\backslash\{x\}$ remains unforced by the values the function takes on B. (The second inequality may undercount $|P(1)|$ since there can be multiple choices for the subset S_x, and there may be a more optimal than greedy way to choose elements from $B \backslash N(N(S_x))$.) $\square$

This bound, though it uses more structural information than the lower bound in (98), may not give a useful result for many posets. Notice that in order for $\sum_{k=2}^{s_x} \binom{s_x}{k} > 0$ there must be at least two points in $N(x)$ whose intersection consists of $\{x\}$ alone. In order for this bound to give useful information, there must be a significant number of vertices in B whose co-degree is relatively small. The whole sum is zero (hence giving a very bad bound!) when vertices in the bottom have high neighborhood overlap, as is the case for $Dk_{n,d}$, and the d-regular circulant graph. However, in both of these cases, the high neighborhood overlap allows us to show that there is a subset of Cherry-free functions of size at least $2^{n(1+\frac{1}{d})}$ which have exactly 0 points forced in A. We can see that the interplay between the sizes of $P(0)$ and $P(1)$ is determined by co-degree.

In the remainder of this section, we show that we cannot get a substantially better upper bound in the case where our poset has co-degree 1, because the bound in (95) remains tight on this class of graphs up to logarithmic factors. We introduce a family of co-degree 1 posets inspired by geometric intuition.

Definition 6.4.4 *The* ***finite projective plane*** *of order d is defined as a set of $m := d^2 - d + 1$ points and m lines. It has the properties that*

- *Given any two distinct points there is exactly one line which passes through both,*
- *Any two lines determine a single point,*
- *Every line contains d points,*
- *d lines pass through each point.*

Let T_d be a bipartite poset with $n = (d^2 - d + 1)$ vertices on the top and bottom. The top set consists of vertices we will call 'lines' and the bottom set contains vertices we will call 'points.' A point adjacent to a line if the line passes through the point. We note that the criteria in the projective plane definition creates a d regular poset with co-degree 1.

We look at the lower bound given in Lemma 6.4.2. $T_d(1)$, the partition class of cherry functions of T_d which contain exactly one fixed point has size at least :

$$|T_d(1)| \geq n\left(1 + \frac{d}{n} + O\left(\frac{\log n}{n}\right)\right).$$

Using the fact that $n = d^2 - d + 1$, we see that

$$n\left(1 + \frac{1}{d} + O\left(\frac{\log(d)}{d^2}\right)\right) \leq \mathfrak{C}(T_d) \leq n\left(1 + \frac{1}{d}\log(d+1)\right). \tag{99}$$

The projective plane was the inspiration for this section. $DK_{N,d}$ is the d-regular graph on N with the highest co-degree and with the largest number of components. It contains an asymptotically maximal number of cherry functions. We initially began looking at the Projective Plane in hopes that because it contains only one component and and has low co-degree that it would give us fewer cherry functions. However, this is how we discovered the interplay between cherry functions with 0 or 1 fixed points exploited in Lemma 6.4.2.

6.5 Horn functions in the Boolean Lattice

Definition 6.5.1 *On a poset P, the* ***supremum*** *of two elements x and y, denoted $x \vee y$ satisfies $x \preceq x \vee y$, $y \preceq x \vee y$ and if both $x \preceq w$ and $y \preceq w$ then $x \vee y \preceq w$. Similarly we define the* ***infimum*** *of two elements, and use the notation $\wedge$.*

Definition 6.5.2 *A poset, L is a* ***lattice*** *for all elements $x, y \in P$ both $x \vee y$ and $x \wedge y$ exist.*

For $x, y \in P$ where P is a lattice, a Horn function f satisfies that $f(x) \wedge f(y) \leq f(x \vee y)$. When the poset in question is the Boolean Lattice, $\mathfrak{B}_n$, we refer to these functions as Boolean Horn Functions. Counting Boolean Horn functions is of interest because these Horn functions are in a 1-1 correspondence with both union closed and intersection-free set systems. If we consider an n-dimensional binary Horn function, f, it defines a union closed family of subsets of the Boolean lattice $\mathfrak{B}_n$. A Boolean function defined on the power set of $[n]$ can be thought of defining a family of sets, those who belong to the family are assigned a 1 by the Boolean function. The property that $f(X) \wedge f(Y) \leq f(X \vee Y)$ enforces exactly that $X \vee Y$, the smallest set containing both X and Y, namely $X \cup Y$, is assigned a 1 whenever both X and Y are. So the number of Horn functions on $\mathfrak{B}_n$ is equivalent to the number of union closed families in $2^{[n]}$. This view of Horn functions will often be useful.

The history of counting Boolean Horn functions is discussed in Pippenger [36]. Burosch *et al.* [10] gave the result that:

$$\binom{n}{\lfloor \frac{n}{2} \rfloor} \leq \log(\alpha(n)) \leq 2\sqrt{2}\binom{n}{\lfloor \frac{n}{2} \rfloor}\left(1 + O\left(\frac{\log(n)}{n}\right)\right) \tag{100}$$

Later, an improved upper bound was given by Alekseyev [1]:

$$\log(\alpha(n)) = \binom{n}{\lfloor \frac{n}{2} \rfloor}\left(1 + O\left(\frac{\log(n)}{n^{\frac{1}{4}}}\right)\right) \tag{101}$$

Pippenger provides a weaker upper bound with a simpler proof which uses entropy of a random variable as an enumeration technique [36]. He proves that:

$$\log(\alpha(n)) = \binom{n}{\lfloor \frac{n}{2} \rfloor}\left(1 + O\left(\frac{\log(n)^{\frac{3}{2}})}{n^{\frac{1}{4}}}\right)\right) \tag{102}$$

Let $\alpha(n)$ represent the number of Horn functions on the Boolean lattice $\mathfrak{B}_n$. The size of the middle layer of $\mathfrak{B}_n$ provides a lower bound for $\log(\alpha(n))$ This follows by construction: If we set f to be one on any subset of this middle layer, we can form a valid Horn function by setting f on the bottom half of $\mathfrak{B}_n$ to 0 and on the top half to 1. This shows there are at least $2^{\binom{n}{\lfloor \frac{n}{2} \rfloor}}$ Horn functions.

We have thus far only defined a cherry function on a two level poset. We can expand this definition to apply to regular posets by calling g a cherry function if on the induced poset formed by any two consecutive levels, g restricted to those levels satisfies the cherry property.

If we consider a Horn function f restricted to two consecutive levels of the Boolean lattice, it is a cherry function since the only time we have a cherry is in the case that $Z = X \cup Y$, and in this case, we will have that $f(X) \wedge f(Y) \leq f(X \vee Y) = f(X \cup Y) = f(Z)$. We can see the former fact with a simple size argument. If X, Y are on level k and Z is on level $k+1$, we have that $|X \cup Y| = |Z| = k+1$ and that $|X \cup Y| \subseteq Z$, so it must be that $Z = X \cup Y$.

Note that even if a function c has the cherry property when restricted to any two level subposet formed by consecutive levels of $\mathfrak{B}_n$, c may not be a Horn function (see

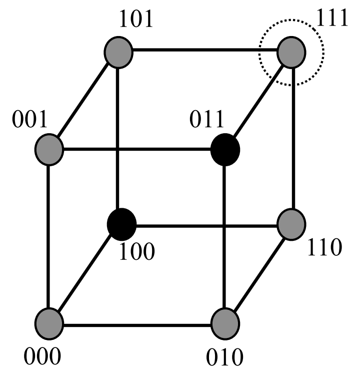

Figure 10: An example in $\mathfrak{B}_3$; cherry function c(shown so black vertices have $c(v) = 1$); circled vertex is a witness that c is not Horn.

figure 6.5). Taking a quick example on $\mathfrak{B}_3$, if $c(100) = c(011) = 1$ and c takes the value 0 elsewhere; This c is a cherry function, which fails to be Horn, since $c(100 \vee 011) = c(111) = 0 < c(100) \wedge c(011) = 1$, violating the Horn condition. Therefore, for $n \geq 2$, $\{\text{Horn functions}\} \subset \{\text{Cherry functions}\}$.

The two level results on the number of cherry functions on bi-regular posets give some information about the number of Horn functions on $\mathfrak{B}_n$, as we can draw some conclusions by applying our results on the bi-regular poset formed when we look at the middle two levels of $\mathfrak{B}_n$. Note that the middle two layers of $\mathfrak{B}_n$ form a poset which has co-degree 1.

A lower bound on the number of cherry functions on the middle two layers is a lower bound on the number of Horn function on all of $\mathfrak{B}_n$ because given any cherry function on the middle two layers, we can extend it to a Boolean Horn function by taking the top half of $\mathfrak{B}_n$ and assigning it function value 1, and assigning everything below the middle two layers the function value 0.

The lower bound from (98) gives that there are at least $2^{\binom{n}{\lfloor \frac{n}{2} \rfloor}\left(1+O\left(\frac{1}{n^2}\right)\right)}$ Horn functions created this way. However, using the structure of $\mathfrak{B}_n$ and a number theoretical argument, a better bound is known.

Theorem 6.5.3

$$\log(\mathfrak{H}(\mathfrak{B}_n)) \geq \binom{n}{\lfloor \frac{n}{2} \rfloor} \left(1 + \frac{1}{n} + o(1)\right).$$

This proof of this bound follows from a proof which appears in [21]. We give this proof in Section 2.5

CHAPTER VII

CONCLUSION

Though the problems of this thesis may seem disparate in both origin and intent, there are two common threads which bind the results. Clearly, one theme is using entropy as an enumeration technique. The other theme stems from the initial problem: How many union-free families are there in the Boolean Lattice? This question seems quite difficult, and remains unanswered to the level of accuracy we desire.

In this concluding section, we would like to draw connections between the preceding chapters, and share how the inspiration for much of the research contained herein is drawn from the quest to count the number of union-free families in $\mathfrak{B}_n$. In our endeavor to enumerate union-free families, we became familiar with techniques that others have used to answer this, or similar, questions. We took each of these techniques, explored their limitations, looked for improvements, and found contexts where similar tools would make progress on related problems.

We first became intrigued with the problem of counting union-free families while looking at a 1987 paper of Pippenger [36]. In that paper, Pippenger gives entropy based proofs for two enumeration problems, counting the number of antichains and union-free families there are in the Boolean Lattice.

Pippenger's result enumerating antichains does not compete in accuracy to the best known results of the same type; at the logarithmic level though he shows matching first order terms, his second order term is off by a factor of $\left(\frac{\log n}{n}\right)^{\frac{1}{2}}$ from the earlier result of Kleitman [29]. However, his use of entropy as an enumeration technique gave his proof the elegance and simplicity that earlier proofs lacked. Indeed, though the results of Korshunov [31] and Alekseyev [1] give asymptotics for the number of antichains itself (as opposed to the logarithm of the number), their proofs are so long and difficult to read that some researchers in the mathematics community doubted their case-based proofs were complete (though it seems that most believed the result itself).

Additional progress was made, in the way of providing simpler proofs on counting antichains in 2001, when Kahn [25] gave an entropic proof whose result rivaled that of Kleitman-Markowsky in accuracy. Kahn was able to use Shearer's lemma, an important entropy tool which has been used effectively in several results since its first use in [11]. Along with an elegant induction proof and a crafty construction, Kahn uses Shearer's Lemma to give his improved result. Kahn's paper basically settled the asymptotic question of antichain enumeration using entropic techniques, However, in light of the success of Kahn's approach, we believed that by using the stronger entropy results available, we would be able to make a similar improvement to Pippenger's approach to union-free families. Thus began our quest.

One thing that holds Pippenger's approach to Dedekind's Problem apart is that he concentrates on using chains as a way to get a handle on antichains instead of working with antichains directly. A single chain can pass through a given antichain at most once, so we gain information from this interaction. This technique provided a way to look at union-free families as well, though the interaction between union-free families and chains is a little more complicated, since the height (hence the intersection with a chain) of a connected component is not limited to a single level. Another complicating factor is that being union-free is an induced property. In an antichain $\mathcal{A}$, if a set A has one neighbor which is in $\mathcal{A}$, it cannot be included. However, if we are looking at a family $\mathcal{F}$ which is union-free, and we are considering a set F to see if it can belong to $\mathcal{F}$, we have to see if there are any sets C, D so that $C \cup D = F$. We can have that $C \cup D \subset F$ or that there exists a set E so that $C \cup D \cup E = F$ (as long as no pair of them makes up all of F). This property is much more difficult to check, hence it is harder to count the number of families which satisfy this property.

In our investigation of the literature about union-free families, we found no lower bound specifically for the largest union-free family. This is a more fundamental problem than how many union-free families there are, so deserved separate inquiry. Knowing the size of a large union-free family is often the first step in enumerating the total number of union-free families; Since the union-free property is hereditary, any subset of the largest union-free family is itself union-free. Finding a large family gives us a useful lower bound. All of the previous literature looking at union-free families used a folklore lower bound (which appeared in [21]) which gives a constructive lower bound for the non-induced version of the union-free property.

As we noted in Chapter 6, where we looked at a loosening of union-free applied to bipartite two-level posets, looking at a property restricted to bipartite posets and then extending the result to ranked partially ordered sets has been a successful tactic for counting the number of several other graph structures: graph homorphisms [20], linear extensions [8], and antichains [26]. Though we have not been able to extend our results to ranked posets, we give results in the two level case with the hopes of finding the proper way to extend them.

The contents of Chapter 5, Dedekind Type Problems, on the surface may not seem related to counting the number of union-free families. However, the motivation for investigating the problems of this section stems from this initial problem. Since Pippenger [36] is able to use similar techniques to count antichains and union-free functions, we studied other techniques which historically have been used to count antichains in hopes of getting better bounds for the number of union-free families. In our efforts to truly understand these techniques we find contexts in which they can be pushed farther than they originally were. We generalize Dedekind's problem in two different ways: first by counting the number of antichains in two posets which are related to the Boolean lattice, and secondly by examining the number of families in $\mathfrak{B}_n$ which forbid an inclusion based structure more complex than antichains.

7.1 Directions for Further Research

Our main motivating problem of counting the number of union-free families in $\mathfrak{B}_n$ is still both open and intriguing. We have concentrated on finding bounds for the logarithm of the number of union-free families which are accurate to the second order term; even at this level of accuracy, the problem remains difficult. We believe that there are several directions which might yield progress. Kahn [26] needed to use a weighted partition function to extend his two-level antichain result to ranked posets. We may in fact need something stronger to fully utilize the information from multiple neighbors. A result like the fractional subadditivity results of Madiman and Tetali [33] may give us this extra strength.

The lower bound in the two level result for the number of cherry functions on a two level poset is currently unsatisfying. We would like to find either an improved bound or a poset for which the greedy selection method we describe matches the lower bound. There is a large gap in the second order terms of the bound, yet all of the posets for which we can calculate the number of cherry functions fall much closer to the upper bound. We would like to explore the expected number of cherry-free functions in a random d-regular two level poset, but at this point, the lack of structure within the extended neighborhoods makes this exploration difficult.

Many of the problems in this thesis (including those in Chapters 5 and 3) compute combinatorial quantities of interest at the logarithmic level, often giving bounds accurate only to first or second order term. This is an obvious place to improve the results of this thesis. The bound in Chapter 5 on the number of antichains in $[t]^n$ is a clear example of this. As the result is stated, the bound holds for all ranges or t and n, but is only a strong bound when t is a constant or growing slowly as a function of n. Clearly there is room to strengthen the result in different ranges of n and t.

A question recently posed by Griggs is 'What is the size of the largest subset of $\mathfrak{B}_n$ containing no diamonds, i.e. four elements a, b, c, d, with $a < b < d$ and $a < c < d$? (Note that b and c need not be incomparable.) In the notation of Chapter 2, letting $\mathcal{D}$ be the set of all diamonds, this questions asks, 'what is $\mathrm{La}(n, \mathcal{D})$?'. Since the middle two layers of $\mathfrak{B}_n$ are obviously a diamond-free family, the natural conjecture is that $\mathrm{La}(n, \mathcal{D}) = \binom{n}{\lfloor \frac{n}{2} \rfloor}(2 + o(1))$. However, the best known result is that $\mathrm{La}(n, \mathcal{D}) \leq 2.5\binom{n}{\lfloor \frac{n}{2} \rfloor}$. Both this question, and the relaxation which requires the sets b and c to be incomparable may yield to the techniques developed in Chapter 2. Of course, there are many interesting small families $\mathcal{P}$ of sets for which these techniques may yield better or new results for $\mathrm{La}(n, \mathcal{P})$.

REFERENCES

[1] ALEKSEYEV, "O chisle semeistv pdomnozhestv," *Diskret. Math*, vol. 1, pp. 129–136, 1989.

[2] ALON, N. and SPENCER, J. H., *The Probabilistic Method.* New York: Wiley-Interscience [John Wiley & Sons], 2 ed., 2000. Wiley-Interscience Series in Discrete Mathematics and Optimization.

[3] BOLLOBÁS, B., *Modern Graph Theory*, vol. 184. New York: Springer-Verlag, 1998. Graduate Texts in Mathematics.

[4] BÓNA, M., *Introduction to Enumerative Combinatorics.* The Walter Rudin Student Series in Advanced Mathematics, McGraw Hill, 2007.

[5] BONDAREVA, O. N., "Some applications of the methods of linear programming to the theory of cooperative games (in russian)," *Problemy Kibernetiki*, vol. 10, pp. 119–139, 1963.

[6] BRÉGMAN, L. M., "Certain properties of nonnegative matrices and their permanents," *Dokl. Akad. Nauk SSSR*, vol. 211, pp. 27–30, 1973.

[7] BRIGHTWELL, G. *Personal Communication*, 2006.

[8] BRIGHTWELL, G. R. and TETALI, P., "The number of linear extensions of the Boolean lattice," *Order*, vol. 20, no. 4, pp. 333–345, 2004.

[9] BUBLEY, R. and DYER, M., "Faster random generation of linear extensions," *Discrete Math*, vol. 201, pp. 81–88, 1999.

[10] BUROSCH, G., DEMETROVICS, J., KATONA, G. O. H., KLEITMAN, D. J., and SAPOZHENKO, A. A., "On the number of Closure operations," *Combinatorics, Paul Erdős is Eighty, Vol. 1*, pp. 91–105, 1993. Bolyai Soc. Math. Stud.

[11] CHUNG, F. R. K., GRAHAM, R. L., FRANKL, P., and SHEARER, J. B., "Some intersection theorems for ordered sets and graphs," *J. Combin. Theory Ser. A*, vol. 43, no. 1, pp. 23–37, 1986.

[12] DE BONIS, A. and KATONA, G. O. H., "Largest families without an r-fork," *Order*, vol. 24, pp. 181–191, 2007.

[13] DE BONIS, A., KATONA, G. O. H., and SWANEPOEL, K. J., "Largest family without $A \cup B \subseteq C \cap D$," *J. Combin. Theory Ser. A*, vol. 111, no. 2, pp. 331–336, 2005.

[14] DEDEKIND, R., *Ueber Gruppen, deren sämmtliche Theiler Normaltheiler sind*, vol. 48. 1897. Math. Ann.

[15] DIESTEL, R., *Graph Theory*, vol. 173. Berlin: Springer-Verlag, 3^{rd} ed., 2005. Graduate Texts in Mathematics.

[16] Engel, K., *Sperner Theory*, vol. 65 of *Encyclopedia of Mathematics and its Applications.* New York, NY: Cambridge University Press, 1997.

[17] Erdős, P., "On a lemma of Littlewood and Offord," *Bull. Amer. Math. Soc.*, vol. 51, pp. 898–902, 1945.

[18] Friedgut, E., "Hypergraphs, entropy, and inequalities," *Amer. Math. Monthly*, vol. 111, no. 9, pp. 749–760, 2004.

[19] Friedland, S., Krop, E., Lundow, P., and Markstrom, K., "Computational validations of the asymptotic matching conjectures," *Preprint*, 2005.

[20] Galvin, D. and Tetali, P., "On weighted graph homomorphisms," *Graphs, morphisms and statistical physics*, vol. 63, pp. 97–104, 2004. DIMACS Ser. Discrete Math. Theoret. Comput. Sci.

[21] Graham, R. L. and Sloane, N., "Lower bounds for constant weight codes," *IEEE IT*, vol. 26, pp. 37–43, 1980.

[22] Griggs, J. and Katona, G. O. H., "No four subsets forming an N," *J. Combin. Theory A*, *to appear.*

[23] Heilmann, O. and Lieb, E., "Theory of monomer-dimer systems," *Comm. Math. Phys.*, vol. 25, pp. 190–232, 1972.

[24] Janson, S., Luczak, T., and Rucinski, A., *Random Graphs.* Wiley Series in Discrete Mathematics and Optimization, New York: John Wiley & Sons, 2000.

[25] Kahn, J., "An entropy approach to the hard-core model on bipartite graphs," *Combin. Probab. Comput.*, vol. 10, no. 3, pp. 219–237, 2001.

[26] Kahn, J., "Entropy, independent sets and antichains: a new approach to Dedekind's problem," *Proc. Amer. Math. Soc.*, vol. 130, no. 2, pp. 371–378 (electronic), 2002.

[27] Karzanov, A. and Khachiyan, L., "On the conductance of order Markov chains," *Order*, vol. 8, no. 1, pp. 7–15, 1991.

[28] Katona, G. O. H. and Tarján, T. G., "Extremal problems with excluded subgraphs in the n-cube," *Graph theory (Lagów, 1981)*, vol. 1018, pp. 84–93, 1983. á Lecture Notes in Math.

[29] Kleitman, D. and Markowsky, G., "On Dedekind's problem: The number of isotone Boolean functions II," *Trans. Amer. Math. Soc.*, vol. 213, pp. 373–390, 1975.

[30] Kleitman, D., "On Dedekind's problem: The number of monotone Boolean functions," *Proc. Amer. Math. Soc.*, vol. 21, pp. 677–682, 1969.

[31] Korshunov, A. D., "O chise monotonnykh bulevykh funktsii," *Probl. Kibern.*, vol. 38, pp. 5–108, 1980.

[32] Lint, J. H. v., *Introduction to Coding Theory.* Graduate Texts in Mathematics, Springer-Verlag, 3rd rev ed. ed., 1999.

[33] Madiman, M. and Tetali, P., "Information inequalities for joint distributions, with interpretations and applications," *IEEE Trans. on Info. Theory*, submitted 2007.

[34] Mattner, L. and Roos, B., "Maximal probabilities of convolution powers of discrete uniform distributions," *Preprint*, 2007.

[35] Ordentlich, E. and Roth, R. M., "Independent sets in regular hypergraphs and multidimensional runlength-limited constraints," *SIAM J. Discrete Math.*, vol. 17, no. 4, pp. 615–623 (electronic), 2004.

[36] Pippenger, N., "Entropy and enumeration of Boolean functions," *IEEE Trans. Inform. Theory*, vol. 45, no. 6, pp. 2096–2100, 1999.

[37] Radhakrishnan, J., "Entropy and counting," *IIT Kharagpur on Computational Mathematics*, vol. Golden Jubilee Volume, 2001.

[38] Riordan, J., *Combinatorial Identities.* R. E. Krieger Pub. Co., 1979.

[39] Sapozhenko, A. A., "The number of antichains in ranked partially ordered sets," *Diskret. Mat.*, vol. 1, no. 1, pp. 74–93, 1989.

[40] Sha, J. C. and Kleitman, D. J., "The number of linear extensions of subset ordering," *Discrete Math.*, vol. 63, no. 2-3, pp. 271–278, 1987.

[41] Shapley, L., "On balanaced sets and cores," *Naval Research Logistics Quarterly*, vol. 14, pp. 453–560, 1967.

[42] Shapley, L., "Cores of conves games," *International Journal of Game Theory*, vol. 1, no. 1, pp. 11–26, 1971.

[43] Shepp, L. A., "The FKG inequality and some monotonicity properties of partial orders," *SIAM J. Algebraic Discrete Methods*, vol. 1, no. 3, pp. 295–299, 1980.

[44] Shepp, L. A., "The XYZ conjecture and the FKG inequality," *Ann. Probab.*, vol. 10, no. 3, pp. 824–827, 1982.

[45] Sloane, N. J. A., "The on-line encyclopedia of integer sequences," `http://www.research.att.com/~njas/sequences/A000372`.

[46] Sperner, E., "Ein satz über Untermegen einer endlichen Menge," *Math. Z.*, vol. 27, pp. 544–548., 1928.

[47] Thanh, H. T., "An extremal problem with excluded subposets in the Boolean lattice," *Order*, vol. 15, pp. 51–57, 1998.

[48] Trotter, W. T., *Combinatorics and Partially Ordered Sets: Dimension Theory.* Johns Hopkins Studies in the Mathematical Sciences, Johns Hopkins University Press, 1992.